Hindawi

Research Article

A Deep Learning Fusion Approach to Diagnosis the Polycystic Ovary Syndrome (PCOS)

Abrar Alamoudi [iD],[1] Irfan Ullah Khan [iD],[1] Nida Aslam [iD],[1] Nourah Alqahtani [iD],[2] Hind S. Alsaif [iD],[3] Omran Al Dandan [iD],[3] Mohammed Al Gadeeb [iD],[3] and Ridha Al Bahrani [iD][3]

[1]Department of Computer Science, College of Computer Science and Information Technology,
 Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia
[2]Department of Obstetrics and Gynecology, College of Medicine, Imam Abdulrahman Bin Faisal University,
 Dammam, Saudi Arabia
[3]Department of Radiology, College of Medicine, Imam Abdulrahman Bin Faisal University, Dammam, Saudi Arabia

Correspondence should be addressed to Abrar Alamoudi; 2200500066@iau.edu.sa

One of the leading causes of female infertility is PCOS, which is a hormonal disorder affecting women of childbearing age. The common symptoms of PCOS include increased acne, irregular period, increase in body hair, and overweight. Early diagnosis of PCOS is essential to manage the symptoms and reduce the associated health risks. Nonetheless, the diagnosis is based on Rotterdam criteria, including a high level of androgen hormones, ovulation failure, and polycystic ovaries on the ultrasound image (PCOM). At present, doctors and radiologists manually perform PCOM detection using ovary ultrasound by counting the number of follicles and determining their volume in the ovaries, which is one of the challenging PCOS diagnostic criteria. Moreover, such physicians require more tests and checks for biochemical/clinical signs in addition to the patient's symptoms in order to decide the PCOS diagnosis. Furthermore, clinicians do not utilize a single diagnostic test or specific method to examine patients. This paper introduces the data set that includes the ultrasound image of the ovary with clinical data related to the patient that has been classified as PCOS and non-PCOS. Next, we proposed a deep learning model that can diagnose the PCOM based on the ultrasound image, which achieved 84.81% accuracy using the Inception model. Then, we proposed a fusion model that includes the ultrasound image with clinical data to diagnose the patient if they have PCOS or not. The best model that has been developed achieved 82.46% accuracy by extracting the image features using MobileNet architecture and combine with clinical features.

1. Introduction

Polycystic ovary syndrome (PCOS) is the most prevailing women's health issue that affects 5 to 10% of women during reproductive age [1]. PCOS is an endocrine disorder and a common cause of infertility [2]. Infertility is defined as the failure in the process of releasing the egg from the ovary. There are many reasons for infertility; one of them is the outgrowth of an unusual number and volume of follicles during the ovulation phase, which is considered the first symptom of PCOS [3]. Ovarian follicles are small cysts that contain a fluid found inside a woman's ovary. Signs and symptoms of PCOS include increased acne, excessive body and facial hair, alopecia, overweight, infertility, and irregular or no periods [4]. The cause of the disease is unknown, while some researchers believe that increased androgen production in ovarian theca cells is the fundamental problem [5, 6]. This hormonal disorder has a massive debate on diagnosis criteria, but the clinical validation of PCOS is usually considered the criteria followed at the Rotterdam

workshop in 2003 [7]. Rotterdam criteria is based on three aspects. If two of them occur, then PCOS may be diagnosed. The three aspects are as follows: (1) high level of androgen (i.e., male sex hormone), (2) oligomenorrhea, and (3) polycystic ovaries (PCOM). The ovary ultrasound image is one of the main tools that can be utilized to predict PCOS at the earliest stage. This image of the ovary contains essential information, such as the number, volume, and position of the follicles [4].

Ultrasound has been the most popular imaging modality in the clinical examination of the patient with ovarian pathology. Ultrasound has several advantages over other medical imaging methods such as computed tomography (CT) and Magnetic Resonance Imaging (MRI). The reason behind that is that ultrasound is of low cost, accessible, and safer and provides real-time results. This imaging technique offers a great opportunity to develop a deep learning model for automatic analysis; it makes this examination more objective and accurate in diagnosis. Deep learning is a powerful approach utilized in imaging analysis and computer vision [8–10]. Automatic classification of PCOS in the earliest stage based on the ultrasound images and clinical data helps in disease recognition.

Diagnosis of the PCOS has different criteria and symptoms that require blood tests, ultrasound tests, and high-quality menstrual data. Therefore, because of the variety of symptoms associated with this syndrome and the absence of a single diagnostic test or method used by clinicians to evaluate patients, medical practitioners are compelled to suggest many clinical test results and unneeded radiological imaging. Consequently, PCOS is diagnosed by excluding irrelevant symptoms or test results, owing to the lack of understanding of its complicated pathomechanism. At the same time, early diagnosis and detection of PCOS with the least number of lab tests and imaging procedures are critical because such a condition leads directly to ovarian dysfunction. In turn, it increases the hazard of infertility, abortion, or even gynecological cancer, as well as mental anguish for patients owing to the wastage of time and money [11]. Currently, the diagnosis of the polycystic ovary morphology in the ultrasound images is performed manually. It is based on specialists' accumulated knowledge and experience in recognizing the morphology and characteristics of the ovary ultrasound images. The specialist's decision after examining the same case's images could sometimes be subjective and variable. Furthermore, ultrasound diagnosis accuracy is higher when conducted by professionals in comparison to less experienced doctors. However, specialist examiners are very limited in number [12], specifically in underdeveloped areas. Moreover, a radiologist's estimate of the follicles' numbers and sizes in the ovary images is carried out manually. Thus, detecting PCOM is a time-consuming and challenging task for radiologists because of the various sizes of follicles and their association with veins and tissues. Furthermore, it makes the images to contain artifacts and speckle noise [13]. The uncertainty in diagnosis can have a long-term impact of women's fertility and hormonal unbalance. Furthermore, this manual framework of

diagnosis could increase the mistakes of the examination, which causes inconvenience for the patient. Thus, it is recommended to propose intelligent computer-aided systems that can offer decision support tools to gynecologists. Using a deep learning model to diagnose women's clinical data and ultrasound images if they have a PCOS. It will also help to overcome the barriers of manual examination of ultrasound images and the assessment of patient clinical data.

Recent implementations of machine learning and deep learning in medical ultrasound images have included diverse activities such as classification [14], segmentation [15], and detection [8, 16]. Ultrasound images of various regions of the body have been used in CAD to diagnose different types of illnesses that can threaten human life, such as breast cancer [17], hydronephrosis [18], and prostate cancer [19]. Moreover, many contributions have been carried out by other researchers in order to identify PCOS by using ultrasound images [4, 13, 14, 16, 20–25]. Several machine learning and deep learning models have been implemented to perform ovary ultrasound image analyses for diagnosis systems, such as SVM [24], NB [22], CNN [20, 21, 25], and VGG-16 [16]. However, many studies used other kinds of data rather than ultrasound images to diagnose PCOS, such as clinical data [11, 26, 27] and ultrasound reports in text format [28]. Moreover, we found that U-Net, VGG-16, and GoogLeNet are some of the CNN architectures achieved significant results for various computer vision tasks such as classification and segmentation of images, as shown in [29, 30], and [16]. Consequently, a deep learning approach was adopted to make the training easier and achieve better performance due to the advantage of deep learning not requiring manual extraction of features from the images. The CNN can automatically extract the features during model training. Moreover, it has been observed that most of the previous works ([20, 21, 29–31]) used a private dataset to build their proposed model. However, the study of [25] used an open-source dataset, but the study of [25]contained 3D images which is not suitable for our study. Moreover, most of these studies, such as [4, 22], and [21], used a small-size dataset to build their model. In general, it has been noticed that there is a shortage of diagnosed PCOS by using ultrasound images. Also, we noticed that the clinical and metabolic data have been used in [26, 27], and [11] to diagnose PCOS patients by using machine learning algorithms. Nonetheless, the diagnosis of PCOS based on Rotterdam criteria includes a high level of androgen hormones, failure of ovulation, and polycystic ovaries on the ultrasound image (PCOM). There is a lack of prior research studies that proposed a deep learning-based fusion model that is able to explore the impact of clinical data along with the ultrasound images to diagnose PCOS. In recent medical imaging literature, there has been a tendency toward using both health data and images in a "fusion-model" to solve complicated problems that a single modality cannot solve, such as skin cancer [32], breast cancer, [33] and Alzheimer disease [34]. Therefore, there is still a need to enhance the model toward achieving better accuracy in PCOS diagnosis by using ovary ultrasound images and clinical data. All the

discussed studies contribute to the process of building the proposed model to detect PCOS.

The main objective of this research is to develop a computer-aided diagnosis CAD model for PCOM diagnosis in order to assist a radiologist in classifying ovary ultrasound images with the aim to reduce the false positive rate and increase the accuracy of the model. Moreover, explore the impact of the clinical features in the diagnosis the PCOS with ultrasound images using a deep learning fusion approach to assist the doctor and radiologist in making better clinical decisions. The rest of this paper is organized as follows. Section 2 discusses the methodology that has been proposed in this paper. Section 3 presents the empirical study with the result that has been obtained. Section 4 provides a discussion of the results, and Section 5 illustrates the conclusions.

2. Materials and Methods

The section provides a detailed description of the dataset used in this research to develop and test the diagnosis model. Then, the proposed model used to diagnose the PCOM using the ultrasound images has been discussed. Also, we discussed the fusion techniques that will be used to combine the ultrasound image with clinical data to build the framework to diagnose PCOS.

2.1. Dataset. In this work, we used the dataset that contains ovary ultrasound images and the clinical data collected from King Fahad Hospital of the University (Khobar, Saudi Arabia). In cooperation with the Department of Radiology, four radiologists were assigned to work on this research who reviewed a total of 1250 patient files that include those patients with documented polycystic ovaries (250) and those who were normal or had other problems in their ovaries (1000). For some of the patients, multiple ultrasound scans were available. However, in the study, only the latest scan of the patient was used. Moreover, only images where the ovary is clearly imaged were selected to be part of the study. For categorization, radiologists classified the chosen images into two groups: a normal morphology (non-PCOM) and those showing sonographic morphology of polycystic ovary (PCOM). The polycystic ovary morphology was defined as an ovary containing multiple uniformly sized follicles that are peripherally placed and are below 1 cm in size, as shown in Figure 1. The images dataset consists of 391 images, 127 PCOM, and 264 normal ovaries non-PCOM.

The process of collecting the clinical data started after completing the image collection to study the impact of the clinical data on the diagnosis of PCOS. This process aims to extract the clinical information from the hospital system for the patients whose ultrasound image has been collected in the previous step. The features were selected with the help of expert opinions and by taking into account recent studies that identified the significance of those attributes on PCOS diagnosis in some manner [11, 22]. During the clinical data collection process for 391 patients whose ultrasound images were already available, we found there are a lot of missing data that require further filtration. Eventually, a dataset contains 22 features and 285 samples. The clinical dataset is composed of a total of 129 PCOS cases and 156 non-PCOS cases. Patients' diagnoses in the dataset are based on laboratory tests, doctor notes, and the radiologist examination of the ultrasound images to PCOS and non-PCOS. The features can be categorized into four types such as demographic (age and marital status), vital signs (BMI, weight, and height), laboratory tests (Follicle Stimulating Hormone (FSH) and Testosterone), and doctor notes such as cycle regularity as shown in Table 1. The dataset contains two data types, four features are nominal/categorical, and 17 features are numeric. Table 1 shows the description of each attribute's features, data type, and possible values.

2.2. Prediction Model. This section presents a detailed description of the proposed techniques that have been used in this study.

(i) The first approach used ultrasound images to diagnose the patient if it has polycystic ovarian morphology (PCOM) or not. This approach starts by reading the images and applying some preprocessing techniques to increase the data quality, such as normalizing the pixel intensities to the range [0, 1], applying the adaptive histogram equalization, and expanding samples by using the image data augmentation technique. Then, using deep learning architectures to extract the image features and classify the patient (VGG-16, VGG-19, InceptionV3, DenseNet121, DenseNet201, and MobileNet), as shown in Figure 2.

(ii) The second approach studies the impact of the clinical data with ultrasound images to diagnose PCOS using the data fusion technique with deep learning models. To prepare the image and the clinical data for the fusion model, preprocessing techniques are applied that have been performed in the previous model when using only the image. On the other hand, the clinical data required some preprocessing methods to prepare the dataset, such as handling categorical data, feature scaling, and dealing with missing data. Two models of joint fusion approaches have been developed and evaluated:

(a) The first is a joint fusion type II, which fuses the clinical features after preprocessing as it is with extracted features from the images; these features were extracted using the deep learning models. Different CNN architectures will be compared to find the most suitable model. The joining features from two modalities will be fed into a feedforward neural network (classification part) to give the final diagnosis, as shown in Figure 3(a).

(b) The joint fusion type I join the learned clinical features with learned imaged features. The features will be learned by applying some dense layers before joining the features from each source. These learned features from the clinical images will be fed to the final model to apply the
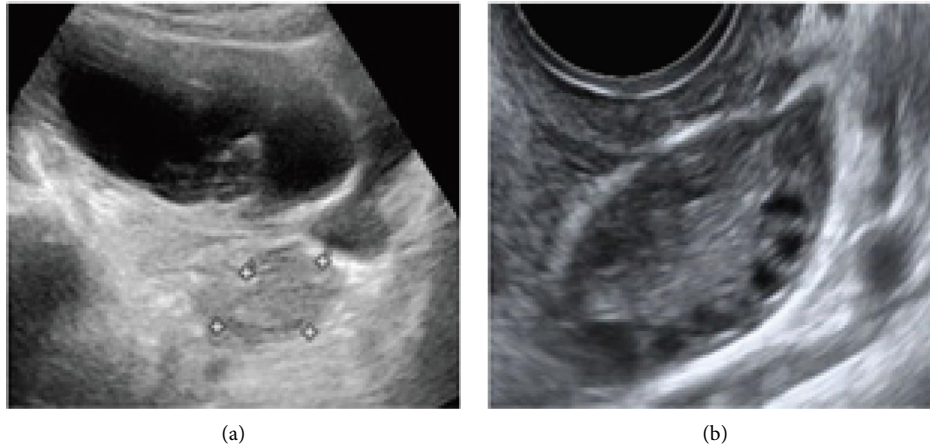
FIGURE 1: Clinical view of ovary ultrasound image: (a) ovary with PCOM and (b) normal ovary.

classification task to diagnose if the patient has PCOS or not, as shown in Figure 3(b).

The proposed techniques are divided into subsections, the first for the transfer learning models and other related techniques such as fusion.

2.3. Transfer Learning. A suitable dataset is essential for the effective operation of any artificial intelligence framework. Data collection and annotation may be difficult, especially when dealing with medical problems. As a result, many issues do not have big datasets that may be employed in deep learning models. As we can notice, the PCOM dataset contains only 391 ultrasound images of normal and polycystic ovaries. In this context, the idea of transfer learning (also known as knowledge transfer) arises. Basically, it reuses the CNN's model developed for a specific task and uses the same weights as a starting point for another problem that has a limited number of images [35]. Transfer learning is a deep learning technique where a model is always pretrained and built using the ImageNet dataset [36], which is a dataset with over 14 million samples and then fine-tuned the same model for a different issue. The following subsections describe in detail the transfer learning models for CNN architectures that have been used and evaluated in this study.

VGG: It is one of the CNN architectures proposed in (ILSVRC) competition of computer vision in 2014 by Simonyan and Zisserman [37]. VGG refers to the visual geometry group lab from Oxford University; the main idea of VGG lies in the fact that it gives significant improvement when using the small convolution filter with (3×3) kernel as compared with a large-sized kernel. The most common architecture of VGG was VGG-16 and VGG-19, which contains 16 or 19 layers, respectively, [38]. It includes three parts convolutional, pooling, and fully connected layers. The VGG16 has 13 convolutional layers and three fully connected layers. Also, it has five pooling layers that are distributed after every 2 or 3 convolutional layers. The main variation between these two models (VGG-16 and VGG-19) was

the VGG-19 model has one additional layer after every three convolutional layers. This model can be used directly for the classification task. The VGG approach's benefit is to reduce the number of parameters and achieve faster convergence [37].

Inception v3: Szegedy et al. [39] first introduced the "Inception" microarchitecture. In 2014, the Inception V3 architecture is based on Szegedy et al. [40], who introduced modifications to the inception module to improve ImageNet classification accuracy and reduce the computational cost. Inception networks (GoogLeNet/Inception v1) have been proven more efficient than VGG based on the number of parameters generated by the network and the cost of the memory and other resources. If any changes are made in inception network should be very careful to secure that the computational advantages are not lost. Thus, there is a barrier in the inception architecture to adapt for diverse use cases that causes the uncertainty of the network's efficiency. Several techniques of improving the network have been proposed in an Inception v3 model to loosen the barriers and restrictions for more effortless model adaptability. Factorized convolutions, regularization, dimension reduction, and parallelized computations are among the techniques used [40]. The Inception v3 consists of a 42-layered deep neural network. The Inception v3 network architecture adopts a convolution kernel splitting technique to break huge volume integrals into little convolutions. It is possible to decrease the number of parameters using the splitting approach, which increases the network training speed while the spatial features can be extracted with greater efficiency [41].

MobileNet: It is a CNN architecture that was developed by Howard et al. and other researchers in 2017 for the classification and detection problem [42]. This architecture is beneficial for mobile and embedded vision applications, and it intends to reduce the number of parameters that will be used in training the model. MobileNet architecture is characterized by being

TABLE 1: Clinical data features.

#	Feature type	Features	Value	Description
1	Demographic	Age	12–66 numeric	Patient age
2		Marital status	S: single, M: married U: unknown nominal	Describe a patient's relationship with a significant other
3	Vital signs	Body mass index (BMI)	Normal: 18.5–24.9 numeric	BMI is a number calculated from a patient weight and height
4		Height	Numeric (cm)	Height weight
5		Weight	Numeric (kg)	Patient weight
6	Radiologist examination	Polycystic ovary morphology	Y: yes, N: no, U: unknown nominal	The result of the ultrasound images if it is showing the morphology of polycystic ovary or normal morphology
7	Doctor notes	Cycle regularity	Y: yes, N: no, U: unknown nominal	Indicate if the cycle of the patient might be regular e or not
8	Laboratory tests	Testosterone	Normal: 7.21–79.31 ng/dl numeric	Indicate the presence of biochemical hyperandrogenism
9		Follicle stimulating hormone (FSH)	Normal: 3.03–8.08 mIU/ml numeric	Indicate the responsiveness of the ovaries to stimulation by the pituitary gland
10		Luteinizing hormone (LH)	Normal: 1.80–11.78 mIU/ml numeric	Indicate the responsiveness of the ovaries to stimulation by the pituitary gland
11		Thyroid stimulating hormone (TSH)	Normal: 0.35–4.94 μIU/ml numeric	Indicate the responsiveness of the thyroid gland to the pituitary gland stimulation
12		Prolactin	Normal: 5.18–26.53 ng/ml numeric	Indicate the activity of the pituitary gland which may affect the cycle regularity
13		Fasting glucose	Normal: 70–99	This is the blood sugar level, and it relates to the insulin resistance
14		Hemoglobin (Hgb) (cbc)	Normal: 12–16 gm/dl numeric	Reflects the oxygen carrying capacity. If low may reflect heavy periods
15		Hemoglobin (Hgb) A1c	Normal: 4–6% numeric	Reflects the blood sugar control over the last 12 weeks
16		Ferritin	Normal: 4.63–204.00 ng/dl numeric	Reflects the iron stores in the body
17		Cholesterol	Normal: <200 numeric	
18		Triglycerides (TG)	Normal: 30–150 numeric	
19		Lipid test — High-density lipoprotein (HDL)	Normal: 35–60 numeric	All lipid profile reflects the risk for atherosclerosis and risk of cardiovascular diseases
20		Low density lipoprotein (LDL)	Optimal <100 mg/dl Borderline 130–159 High >160 mg/dl numeric	
21		Risk factor	Normal: 4.4–5.0 Chol/HDL numeric	
22	—	Image ID	ID number (ex: 1, 2, 3) like the reference of the image	Reference for ultrasound image that belongs to the patient
23		Class	PCOS/non-PCOS nominal	Target class

Input : Ultrasound Images

Preprocessing : pixel normalization, data augmentation technique, adaptive histogram equalization

Feature Extraction & Classification : Deep learning models
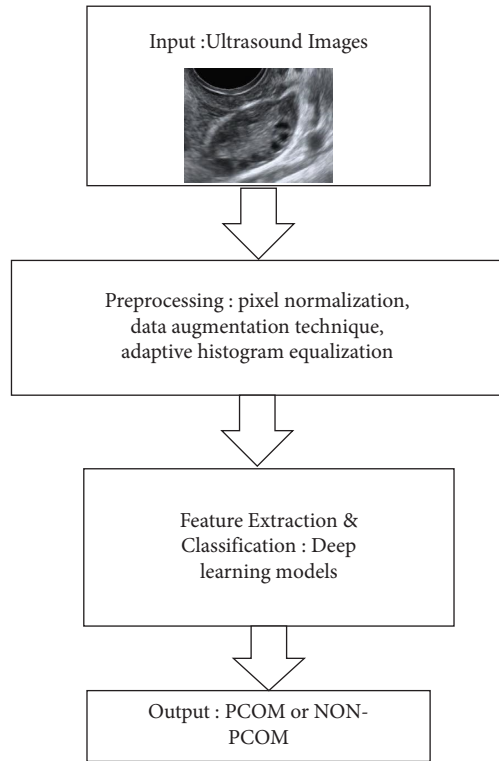
Output : PCOM or NON- PCOM

FIGURE 2: Diagram of the model for the PCOM diagnosis.

streamlined by using the depthwise separable convolutions that help to construct lightweight networks to decrease the model complexity and size. Depthwise separable convolution includes two main operations. The depthwise convolution and pointwise convolution [42]. Depthwise convolution applies a spatial convolution in each channel. Therefore, it has the same number of output channels as input channels. Pointwise convolution: applied the convolution with 1×1 kernel size that combines the output of the depthwise convolution to change the dimension. The MobileNet architecture is made up of 28 layers [43].

DenseNets: This model refers to densely connected convolutional networks, which is a CNN that uses dense block to connect all layers directly with each other to make a dense connection between layers. This architecture motivates to solve the issue of the standard CNNs when the information path starts from the input layer and passes through so many layers to reach the output layer, that it may vanish by the time reaching the end of the network [44]. The barriers can be handled in this architecture when each layer gets new inputs from all succeeding layers and passes on its feature maps to all following layers to retain the feed-forward nature. The first layer connects to the second, third, and fourth layers, and the second layer connects to the third, fourth, fifth, and so on. The DenseNet has $L\ (L+1)/2$ connection with L layers, whereas traditional CNN has an L connection. DenseNet has many advantages, such as needing less number of parameters than traditional CNN due to this type of connection, enhancing feature

distribution, and encouraging feature reuse [44]. DenseNets architectures are split into dense blocks, while the dimensions of the feature map remain constant within a block, but a difference in the number of filters within them. Transition layers reside between the DenseBlocks for down sampling to decrease the number of channels. A batch-normalization layer, 1×1 convolution, and a 2×2 average pooling layer create the transition layers [45]. The DenseNet121 version of DenseNet was pretrained on the ImageNet dataset. Number "121" in the DenseNet121 network refers to the layers with trainable [46]. DenseNet121 is made up of many dense blocks that contain a different number of layers (repetitions) with two convolutions; one has a 1×1 kernel as the bottleneck layer to reduce the feature number and a 3×3 kernel to execute the convolution operation. Also, it has many transition layers between the dense blocks. Therefore, Dense-Net121 has one 7×7 convolution, 58 3×3 convolution, 61 1×1 convolution, four average pooling, and one fully connected layer. Our proposed classifier replaces the final fully connected layer, as well as the SoftMax activation

These architectures have been obtained for feature extraction and classification task by using the model layers, from the input layer to the last pooling layer (front layers) for feature extraction and can use for the rest of the model layers (fully connected layer) for the classification task.

2.4. Fusion Technique. CNNs have achieved significant success in a wide range of applications. Combining multimodal fusion with CNNs seems to be a promising area for future research. Data fusion refers to the process of combining and associating the data and information from many modalities to provide more accurate, consistent, and complete information to develop machine learning models that outperformed the individual data modality [47]. Data fusion approaches have been widely used in multisensor environments to combine and aggregate data from many sensors; however, similar techniques may also be used in other areas, such as text processing. Data fusion in multisensor settings aims to reduce detection error probability and increase dependability by combining data from various dispersed sources [48].

There are many techniques for multimodal data fusion, such as early fusion, joint fusion, and late fusion. Early fusion, also called features level fusion, refers to combining various types of input data into one vector feature for training and then feeding this vector to a single machine learning model, as shown in Figure 4(a). Combining the input can be done using many methods that include concatenation, pooling, or applying a gated unit. Early fusion type I combines the actual features, whereas early fusion type II combines the extracted features whether through manual extraction, image analysis tools, or a learned representation from a neural network [49]. Joint fusion which has been used in this research activity, also called intermediate fusion [50], is the approach of feeding learned feature representations
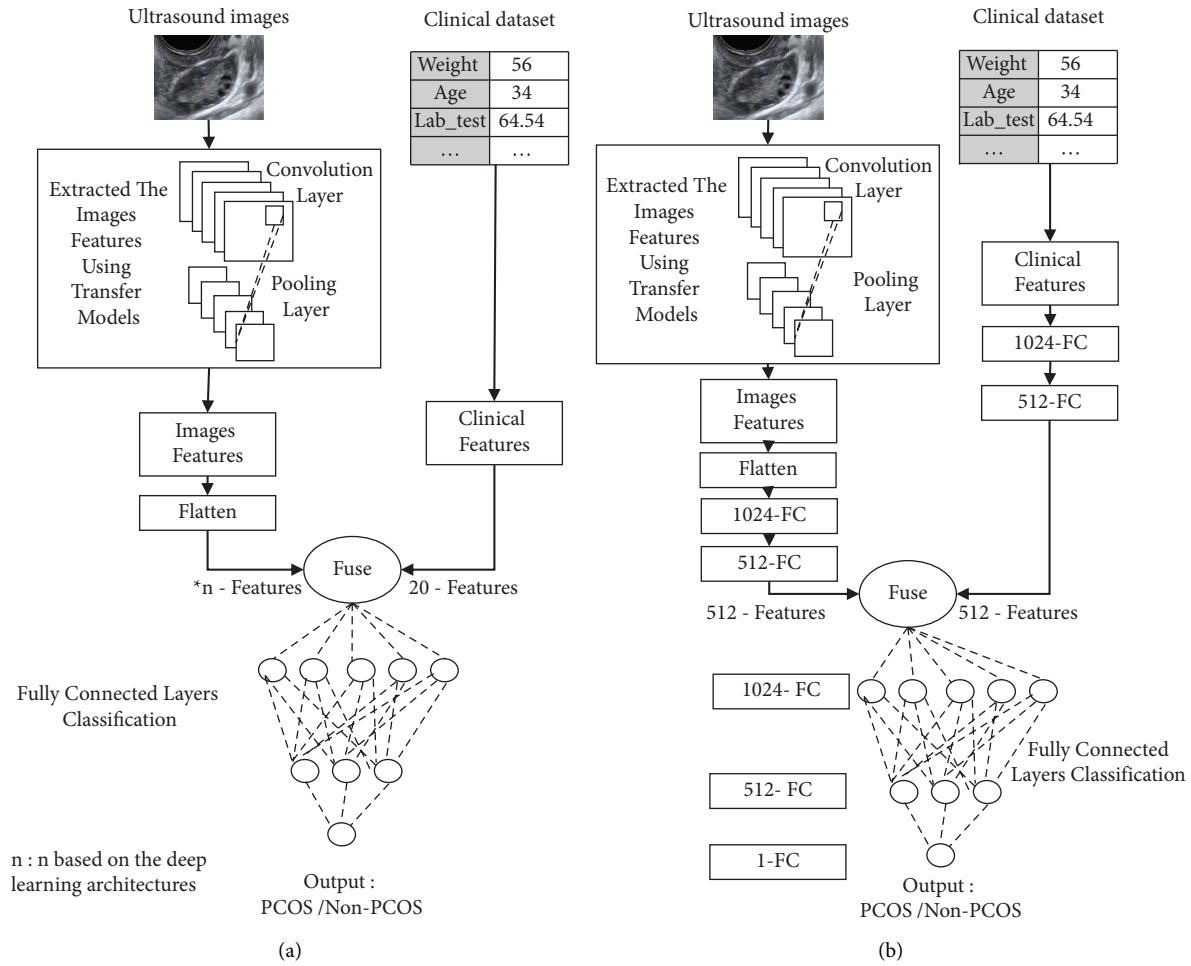
FIGURE 3: Diagram of the model for diagnosis of the PCOS by combining the image with the patient clinical dataset. (a) Joint fusion type II and (b) joint fusion type.

from intermediate layers of neural networks into a final model with features from other modalities, as shown in Figure 4(b). The join fusion was similar to early fusion, except that the loss is fed back to the neural network that is responsible for extracting the feature, which helps to enhance the feature representations in each training reiteration [49]. The joint fusion type I is defined when the input features from each modality are extracted and learned before combining them. However, joint fusion type II does not require the feature extraction step for all input features to be identified as joint fusion.

3. Experiments and Results

3.1. Experiment I: Ultrasound Images. In this part, the experiments to predict PCOM are explained using the ultrasound image dataset. The following part gives a view of the proposed framework, which goes through several phases, starting by applying preprocessing techniques that aim to improve the dataset's quality. Then, applying the feature extraction and classification task using deep learning architectures. The dataset was divided into training and testing sets using an 80:20 holdout to construct and validate the

prediction model. The training set was utilized to learn the model and fine-tune the model's parameters, while the testing set was used to assess the model for hyperparameter tuning and select the most proper model.

3.1.1. Preprocessing. Preprocessing is a critical phase before feeding the data to be processed in the feature extraction stage. This step will provide data with higher quality and allow essential information included in the image to be retrieved more readily. In the preprocessing phase, the data augmentation and AHE techniques were applied.

The preprocessing step starts by loading the images, and each extracted ovarian image is a 224×224-pixel size. Then, the pixel normalization was applied by scaling all pixel intensities to the range [0, 1]. As we have mentioned earlier, since a limited sample size of the current dataset is in hand, there is a need to increase each sample. The image sample expansion has been carried out using the image data augmentation technique. This technique artificially expands the dataset size by creating modified copies of images through random geometric transformations, such as flipping, cropping, rotating, and random erasing. Moreover, data
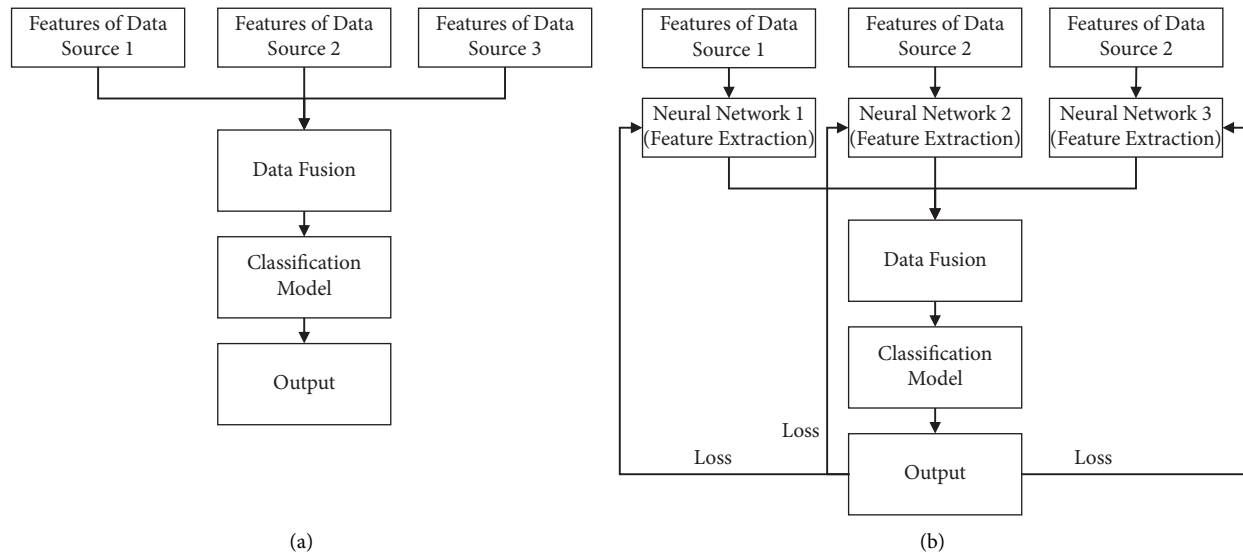
FIGURE 4: Early fusion and joint fusion. (a) Early fusion. (b) Joint fusion.

augmentation is helpful for making the model not see and use the same batch of input during each training iteration. This will help reduce the overfitting problem [51]. The augmented images are generated on the fly while the model is still being trained.

Maheswari et al. [22] have shown the benefits of using AHE for removing noise. Following their recommendations, we applied the AHE for all images before the training phase. The contrast level is raised using an AHE to distinguish the background from the foreground. Also, it is considered a practical approach to reduce the speckle noise for local minima extraction, as shown in Figure 5. These speckles could be causing the reduction of obscures and contrast details more diagnostically [22]. After the image goes through the enhancement process, the rest of the model goes through deep learning architectures.

3.1.2. Applying the Deep Learning Architectures. After obtaining the preprocessing phase, the CNN models may be used to produce a deep network that can properly learn and extract the features in the ovarian ultrasound images. CNN's architectures are applied to predict if the patient has polycystic ovarian morphology or not. Pretrained deep learning models are used to extract the images features, where the features extraction process is carried out in the front layers of the network and then modified the fully connected layers (last few layers in the network) to produce a deep network that can properly learn and diagnose PCOM using ultrasound images. We perform experiments using 6 well-known CNN architecture: VGG16, VGG19, InceptionV3, DenseNet121, DenseNet201, and MobileNet. These architectures are selected to be examined and compared with our dataset based on their performance in previous studies. Moreover, they are the most popular and widely used architectures in healthcare fields [8, 52].

For all pretrained models that have been used in the experiment, we used their convolutional and pooling layers

as it is in the original architecture to extract the image features, but we will not use their fully-connected output layers to perform the predictions task. Instead of that, we construct a new fully connected layer that is suitable for our problem. It means the classification layer trained on the ImageNet dataset was removed. After that, the new fully connected layer will be constructed that is suitable for our problem and append on top of the architecture. During the experiment, many modifications are carried out on the fully connected layer to enhance the performance, such as adding a dropout layer to all hidden layers with a different probability, changing the learning rate value, changing the number of dense layers, and testing different optimization methods. The best fully connected model that has been reached is shown in Figure 6. The models are trained for 100 epochs with a batch size of 32 images. Also to prevent overfitting, the dropout was used with a probability of 0.5; it should also be noted that different values were tested for the dropout and we found the most suitable results when we use 0.5. Also, the initial learning rate value was 0.00001 which controls how much to change the model in response to the estimated error each time the model weights are updated and applied the "Adam" algorithm for optimization. Adam's optimization method computes individual adaptive learning rates for different parameters from estimates of the first and second moments of the gradient to compute individual adaptive learning rates for different parameters [53]. Each model architecture has been assessed using different metrics such as accuracy f-score, precision, recall (sensitivity), and specificity. The code is available upon request.

3.2. Result of Experiment I. This section presents and discusses the results of applying the previously defined experiment to diagnose polycystic ovary in the ultrasound images. The full Deep Learning models are applied for feature extraction and classification. Experiments using 6 well-known CNN architectures are performed: VGG16,
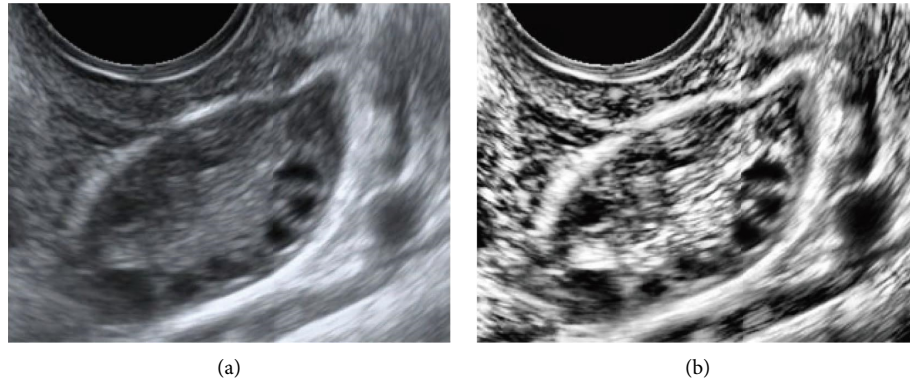
Figure 5: Clinical view of ultrasound ovary image before (a) and after (b) applied adaptive histogram equalization.
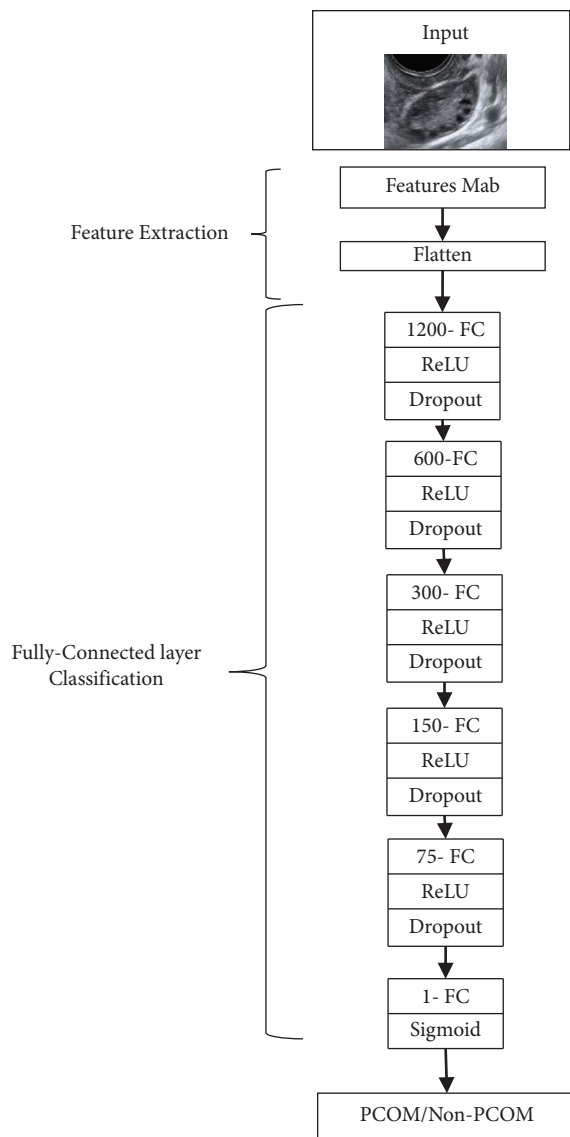

Figure 6: Structure of the fully connected layers in the network architectures.

VGG19, InceptionV3, DenseNet121, DenseNet201, and MobileNet. The performance evaluation on the test set using the deep neural network architectures is presented in Table 2. The Inception model was the best performing model, getting an 84.81% accuracy, 69.57% precision, 72.73% F1-score, 76.19% recall, and 87.93% specificity. Due to uneven class distribution in our dataset, we prefer to focus on the F1-score measure, which is achieved the best result than other models. Also, we noticed that models InceptionResNet, ResNet_152, and EfficientNet_B3 have not achieved a good result with our dataset; accordingly, these architectures are excluded from the following experiment.

Table 3 depicts the confusion matrix for the best accurate model while applying Inception architecture for the feature extraction and classification task. This table shows the network's overall classification rate and accuracy. The number of correctly classified instances is shown in the diagonal cells, whereas the number of misclassified cases is shown in the off-diagonal cells.

This experiment considers only the ultrasound image to diagnose PCOM. Although, it is better to extend the model to include clinical features for accurate diagnosis of PCOS as a syndrome, where images allow only diagnosis of the polycystic morphology.

3.3. Experiment II: Ultrasound Images + Clinical Data. This section discusses the PCOS diagnosis model based on combining the ultrasound images with the clinical dataset to construct the fusion deep learning model. To build the fusion model using a multi-input, we need two branches, one for the clinical dataset and the other for the ultrasound images. The images and clinical dataset are split into training and testing set using 80 : 20 holdout for the 285 samples. Before starting the fusion model, the preprocessing phase is required for the images and clinical data to be ready for the fusion deep learning model. For the images dataset, the preprocessing step starts by loading the ultrasound images with the size of 331×331-pixel then the same preprocessing techniques are applied that have been performed in the first

TABLE 2: The results for all models in applying the feature extraction and classification task using deep learning architectures.

Model	Accuracy (%)	Precision (%)	F1-score (%)	Recall (sensitivity) (%)	Specificity (%)
VGG16	70.89	65.22	56.60	50.00	83.67
VGG_19	67.09	65.22	53.57	45.45	82.61
Inception v3	**84.81**	69.57	**72.73**	**76.19**	**87.93**
MobileNet	69.62	56.52	52.00	48.15	80.77
DenseNet 121	68.35	**73.91**	57.63	47.22	86.05
DenseNet 201	75.95	69.57	62.75	57.14	86.27
ResNet_152	72.15	4.35	8.33	100.00	71.79
EfficientNet_B3	70.89	4.35	8.00	50.00	71.43
InceptionResNet	73.42	13.04	22.22	75.00	73.33

TABLE 3: Confusion matrix for inception model.

		Predicted	
		PCOM	Non-PCOM
Actual	PCOM	**16**	7
	Non-PCOM	5	**51**

model except for data augmentation, which is discussed in Section 3.3.1. Also, the preprocessing phase for the clinical dataset includes many techniques: handling categorical data, feature scaling for the continuous data, and dealing with the missing data.

To evaluate if the fusion of image and clinical features can have an impact on PCOS prediction, we explore and evaluate different fusion approaches to combine image features with clinical features. In the beginning, the late and joint fusion have been compared. The late fusion does not give any promising results because all cases have been diagnosed as negative cases. Then, the empirical study focused on joint fusion by exploring and comparing different stages to fuse features from multimodal data. Two model architectures have been performed and evaluated as follows.

3.3.1. Joint Fusion Type II.
The first approach of the fusion model feeds clinical features as it is in the first branch after applying the preprocessing techniques. The second branch performs the deep learning model over the image data to extract the image features. This branch compares different architectures, including VGG16, VGG19, InceptionV3, DenseNet121, DenseNet201, and MobileNet, to find the best model that is applied to diagnose PCOS. These two branches will be concatenated together to perform the classification part that gives the final diagnosis, as shown in Figure 3(a). The classification part consists of 2 dense layers and a dropout layer with a probability of 0.2.

3.3.2. Joint Fusion Type I.
The second approach applies some dense layers in the first branch to handle the clinical features to give 250 neurons as output for this branch. Also, for the second branch, after image features are extracted using the deep learning model, two dense layers will be performed to have the same number of neurons, which is 250. Then, these learned features from two branches will be combined using

the same classification part in the previous approach to produce a final prediction, as shown in Figure 3(b).

These models are trained for 100 epochs with a batch size of 10. Also, the initial learning rate value is 0.001, and the "Adam" algorithm for optimization is used. The code is available upon request.

3.4. Result of Experiment II.
This section discusses the result that has been obtained while applying the fusion model that aims to combine the ovary ultrasound images with clinical data to diagnose PCOS. Two approaches are proposed for the fusion model. The first one is joint fusion type II which fuses the raw features for the clinical data with image features that are extracted using the deep learning models such as VGG_16, VGG19, InceptionV3, DenseNet121, DenseNet201, and MobileNet. Then, the classification part is applied using fully connected layers. In contrast, the joint fusion type I fuses the learned feature for images and clinical data as input to fully connected layers to make the final diagnosis. Table 4 shows the classification accuracy, precision, F1-score, recall (sensitivity), and specificity for the joint fusion model type II while using different deep learning models to extract the image features. It is noticed in Table 4 that joint fusion model type II, while the image feature extraction using VGG-16 and VGG-19 outperformed the other models in all metrics. The result while using the VGG-16 is 77.19%, 61.54%, 71.11%, 84.21%, and 73.68% for accuracy, precision, F1-score, recall (sensitivity), and specificity, respectively. On the other hand, the model using VGG-19 achieved 75.44%, 80.77%, 75.00%, 70.00%, and 81.48% for accuracy, precision, F1-score, recall (sensitivity), and specificity, respectively. We noticed that the first model is higher based on the accuracy and sensitivity, but the model used VGG-19 is more elevated on the precision, F1-score, and specificity. In medical problems, recall (sensitivity) metrics play an important role due to the cost associated with false-negative will be extremely high, that is because of the complications related to PCOS and the risk of developing health problems in later life will be increased if the PCOS is not diagnosed in the early stage. But F1-score for the second model using VGG-19 is better than the first model using VGG-16 by approximately 4%, which gives a harmonic mean of precision and recall (sensitivity), and it indicates that we have low false positives and low false negatives.

TABLE 4: Performance for all models in the joint fusion type II.

Model	Accuracy (%)	Precision (%)	F1-score (%)	Recall (sensitivity) (%)	Specificity (%)
VGG16	**77.19**	61.54	71.11	**84.21**	73.68
VGG_19	75.44	**80.77**	**75.00**	70.00	**81.48**
Inception v3	73.68	57.69	66.67	78.95	71.05
MobileNet	68.42	53.85	60.87	70.00	67.57
DenseNet 121	61.40	73.08	63.33	55.88	69.57
DenseNet 201	68.42	61.54	64.00	66.67	69.70

TABLE 5: Performance for all models in the joint fusion type I.

Model	Accuracy (%)	Precision (%)	F1-score (%)	Recall (sensitivity) (%)	Specificity (%)
VGG16	71.93	73.08	70.37	67.86	75.86
VGG_19	75.44	73.08	73.08	73.08	77.42
Inception v3	78.95	84.62	78.57	73.33	85.19
MobileNet	**82.46**	**84.62**	**81.48**	78.57	**86.21**
DenseNet 121	70.18	42.31	56.41	**84.62**	65.91
DenseNet 201	77.19	80.77	76.36	72.41	82.14

On the other hand, Table 5 presents the result achieved in the joint fusion model type I with the same evaluation metrics. This model achieves a better result while using the MobileNet model to extract image features with 82.46% accuracy, 84.62% precision, 81.48%% F1-score, 78.57% sensitivity, and 86.21% specificity.

Figure 7 presents a detailed comparison between the best model for joint fusion types I and II. The confusion matrix shows a summary of the prediction results of a classification model. The joint fusion type II has made a total of 57 predictions. Out of 57 predictions, 43 are true predictions and 14 are incorrect predictions. Whereas joint fusion type I has made out of 57 predictions, 47 are true predictions and 10 are incorrect predictions. Based on that, the joint fusion type I outperformed another model, and it is more suitable for our problem to diagnose PCOS using the ovary ultrasound image with clinical data. As noticed, the fusion model for ultrasound images with clinical data to diagnose PCOS gives promising results to develop CAD systems that can assist the doctor in making the right decision.

4. Discussion

In this research, three main experiments are performed to develop the CAD model that is applied to assist the radiologist and gynecologist during the diagnosis of PCOS. In the first experiment, the PCOM is diagnosed based on the ovary's ultrasound image. The proposed model achieved 84.81% accuracy using the Inception model, as shown in Figure 8. This model has achieved good outcomes compared to the result obtained in state-of-the-art studies that invest in detecting the PCOM using the ultrasound images of the ovary. Srivastava et al. [16] achieved 92.11% accuracy using VGG-16, but their work goal is more general and simple because they detect whether an ovarian cyst is present in the image or not, while there are many types of cysts; one of them is PCOS. However, [20, 54] achieved 78.1% and 80.84% accuracy, respectively, whereas the current research results

are higher than these mentioned research. Moreover, [20, 54] performed the classic approach by applying feature extraction using the Gabor wavelet method and then applying a classification task using CNN or Elman neural network. Abdullah et al. [3] got higher accuracy of 93.02% than our work, but they also applied the classical techniques using Gabor wavelet as a feature extractor and a modified backpropagation as a classifier. Also, the authors did not specify the number of images used in the experiment, which is one of the important and influential criteria. While Cahyono et al. and others [21] achieved a 100% micro-average F1-score using the proposed architecture of CNN, which cannot be compared with current work due to using a limited number of samples in the dataset and an unbalanced dataset (total of 40 non-PCO data and 14 PCO).

The second experiment aims to explore the effect of the clinical features in diagnosing PCOS using ultrasound images. A fusion deep learning model is developed that fuses the ultrasound image features with clinical features to produce a final diagnosis of PCOS. The best model that has been developed extracts the image features using MobileNet architecture and joins them with learned clinical features. These combined learned features are fed to fully connected layers to apply the classification task and provide the final diagnosis. The model has achieved 82.46% accuracy, which is less than the result gained while using only image or clinical features separately, as shown in Figure 8. Nevertheless, the fusion model that combined the ultrasound image with clinical data outperformed the model that uses only images based on the other metrics, precision, F1-score, recall (sensitivity), and specificity, which gives a more accurate description of the model's performance. It is also worth noting that the diagnostic process using the ultrasound images only diagnoses the polycystic morphology (PCOM), which is one of the Rotterdam criteria to diagnose PCOS, while the model that included the images and patient information aimed to diagnose the PCOS. Meanwhile, the achieved results represent an advancement towards automated PCOS detection that is

Joint fusion type II – (VGG-19)

Confusion Matrix

		Predicted	
		PCOS	Non-PCOS
Actual	PCOS	21	5
	Non-PCOS	9	22

Joint fusion type I – (Mobilenet)

Confusion Matrix

		Predicted	
		PCOS	Non-PCOS
Actual	PCOS	22	4
	Non-PCOS	6	25

FIGURE 7: Confusion matrix of the best fusion model of types I and II.
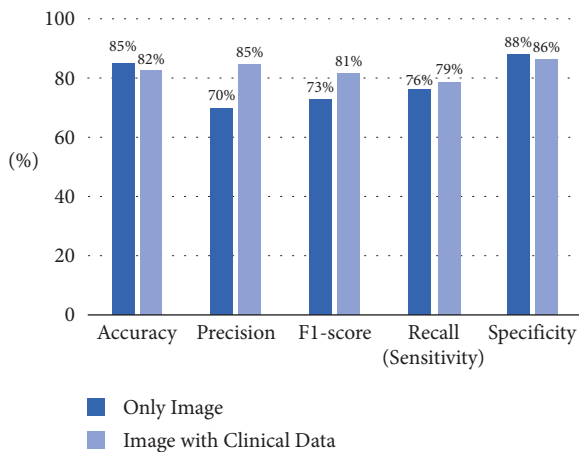

FIGURE 8: The performance of the diagnosis model while considering only image and fusion model.

based on the multimodality fusion model. Most of the state-of-the-art approaches reported in the literature use only clinical images; they do not consider the patient clinical information with images except the study of [22] that used images and clinical features; they achieved 98.63% accuracy. This paper has utilized traditional feature extraction techniques such as modified furious flies and traditional classifiers ANN and NB while using a limited number of samples 68. The development of a fusion model for ovary ultrasound images with clinical features to diagnose PCOS is still a new field and needs further development.

Meanwhile, the findings show that progress toward automated PCOS identification is being made. Also, the results of the fusion model are acceptable; it indicates that the clinical features affect the PCOS diagnosis. However, it requires more improvement that could be attained by including more features that describe the patient's situation like the patient's lifestyle, diet, and some features that appear on the patient, such as dark areas on the skin, acne, pimples, weight gain, and hair loss. However, it is still necessary to perform exhaustive testing and experiments using data from different medical centers before deploying such a system.

5. Conclusion

This research activity aims to propose a PCOS diagnosis and analysis model for CAD systems. Two datasets have been collected; one for ovary ultrasound images and another for clinical data that include vital sign information, lab test results, and some symptoms, which help diagnose PCOS. Multiple deep learning frameworks are proposed to implement the AI-based CAD. The first proposed model diagnoses the PCOS morphology with ovary ultrasound images using the deep learning model for the automated diagnosis of PCOM with the aim to reduce the false positive rate and increase the performance of the proposed model. Many experiments have been performed to achieve this goal using different CNN architectures. The proposed model employs the Inception model fine-tuned with the ultrasound images dataset. Fine-tuning is carried out by modifying the last layers of the Inception network, which is responsible for doing the classification task (fully connected layers). This model can determine whether the ultrasound images show polycystic or not and has obtained 84.81% accuracy, 69.57% precision, 72.73% F1-score, 76.19% sensitivity, and 87.93% specificity. The result achieved gives a notable improvement on a benchmark dataset, which indicates the promising future of obtaining a CAD system that is able to assist the radiologist in classifying ovary ultrasound images. This research activity has presented a study to analyze the impact of combining the image and clinical features using a deep learning model to diagnose PCOS. Two fusion models are compared and analyzed: joint fusion types I and II. The finding of this experiment has shown that the joint fusion type I has outperformed with 82.46% accuracy, 84.62% precision, 81.48%% F1-score, 78.57% sensitivity, and 86.21% specificity. To sum up, this research highlighted the relevance of clinical features in PCOS diagnosis and proved that patient clinical information is necessary for diagnosing PCOS. This automated model can aid the physician in saving the time required to assess patients and reduce the risk associated with delayed PCOS diagnosis.

Some limitations of this study could be addressed in future research. First, the number of samples in the dataset is limited due to the difficulty in the process of collecting the

images and clinical data. The study has also used the dataset from a single center. In addition, one of the main limitations of this research is the lack of available computational resources that have been utilized to perform the empirical study, which in turn impacts the model's performance.

Despite the promising outcomes achieved by the proposed approaches, there are many considerations that might be improved in the future. Increasing the number of samples in the dataset (images and clinical data) helps improve the model performance, generalizes it successfully and reduces errors and misclassification. Although the current study approach only considers one image per patient, it is possible to have many images from various angles for the same patient. It is therefore desirable that additional ultrasound images can be taken into consideration. Also, the clinical dataset could be improved by adding more features that describe some symptoms observed in the patients, including acne, hirsutism, and other signs of hyperandrogenism, amenorrhea, and infertility. Moreover, the deep learning model that has used ultrasound images to diagnose PCOM can be used to classify other types of ovarian cysts, such as functional cysts, endometrioma or endometrioid cysts, dermoid cysts, hemorrhagic ovarian cysts, and PCOS.

Data Availability

The Images and clinical data used to support the findings of this study are restricted by the Ethics Research Committee of Imam Abdulrahman bin Faisal University in order to protect patient privacy.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

[1] L. E. Kenigsberg, C. Agarwal, S. Sin et al., "Clinical utility of magnetic resonance imaging and ultrasonography for diagnosis of polycystic ovary syndrome in adolescent girls," *Fertility and Sterility*, vol. 104, no. 5, pp. 1302–1309, 2015.

[2] T. T. Lee and M. E. Rausch, "Polycystic ovarian syndrome: role of imaging in diagnosis," *RadioGraphics*, vol. 32, no. 6, pp. 1643–1657, 2012.

[3] F. S. Abdullah and N. Jondri, *Modified Backpropagation Algorithm for Polycystic Ovary Syndrome Detection Based on Ultrasound Images*, in *Proceedings of the International Conference on Soft Computing and Data Mining*, Berlin, Germany, December 2017.

[4] C. Gopalakrishnan and M. Iyapparaja, "Active contour with modified Otsu method for automatic detection of polycystic ovary syndrome from ultrasound image of ovary," *Multimedia Tools and Applications*, vol. 79, no. 23–24, pp. 17169–17192, 2020.

[5] R. Azziz, "PCOS: a diagnostic challenge," *Reproductive BioMedicine Online*, vol. 8, no. 6, pp. 644–648, 2004.

[6] V. L. Nelson, R. S. Legro, J. F. Strauss, and J. M. McAllister, "Augmented androgen production is a stable steroidogenic phenotype of propagated theca cells from polycystic ovaries," *Molecular Endocrinology*, vol. 13, no. 6, pp. 946–957, 1999.

[7] The Rotterdam Eshre/Asrm-sponsored Pcos consensus workshop group, "Revised 2003 consensus on diagnostic criteria and long-term health risks related to polycystic ovary syndrome (PCOS)," *Human Reproduction*, vol. 19, no. 1, pp. 41–47, 2004.

[8] S. Liu, Y. Wang, X. Yang et al., "Deep learning in medical ultrasound analysis: a review," *Engineering*, vol. 5, no. 2, pp. 261–275, 2019.

[9] I. U. Khan, N. Aslam, F. M. Anis et al., "Amniotic fluid classification and artificial intelligence: challenges and opportunities," *Sensors*, vol. 22, pp. 4570–4612, 2022.

[10] I. U. Khan, N. Aslam, F. M. Anis et al., "Deep learningbasedcomputer-aided classification of amniotic fluid using ultrasound images from Saudi arabia," *Big Data and Cognitive Computing*, vol. 6, no. 4, p. 107, 2022.

[11] A. Denny, A. Raj, A. Ashok, C. M. Ram, R. George, and I. Hope, "Detection and prediction system for polycystic ovary syndrome (PCOS) using machine learning techniques," *IEEE Region 10 Annual International Conference, Proceedings/ TENCON*, vol. 2019, Article ID 8929674, 678 pages, 2019.

[12] F. Christiansen, E. L. Epstein, E. Smedberg, M. Åkerlund, K. Smith, and E. Epstein, "Ultrasound image analysis using deep neural networks for discriminating between benign and malignant ovarian tumors: comparison with expert subjective assessment," *Ultrasound in Obstetrics and Gynecology*, vol. 57, no. 1, pp. 155–163, 2021.

[13] O. R. Isah, A. D. Usman, and A. M. S. Tekanyi, "A hybrid model of PSO algorithm and artificial neural network for automatic follicle classification," *Int. J. Bioautomation*, vol. 21, no. 1, pp. 43–58, 2017.

[14] A. Nabilah, R. Sigit, T. Harsono, and A. Anwar, "Classification of ovarian cysts on ultrasound images using watershed segmentation and contour analysis," in *Proceedings of the 2020 International Electronics Symposium (IES)*, pp. 513–519, Surabaya, Indonesia, September 2020.

[15] E. Setiawati and A. B. W. Tjokorda, "Particle Swarm Optimization on follicles segmentation to support PCOS detection," in *Proceedings of the 2015 3rd International Conference on Information and Communication Technology (ICoICT)*, pp. 369–374, Nusa Dua, Indonesia, May 2015.

[16] S. Srivastava, P. Kumar, V. Chaudhry, and A. Singh, "Detection of ovarian cyst in ultrasound images using fine-tuned VGG-16 deep learning network," *SN Comput. Sci.*vol. 1, no. 2, pp. 81–88, 2020.

[17] W. K. Moon, Y. W. Lee, H. H. Ke, S. H. Lee, C. S. Huang, and R. F. Chang, "Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks," *Computer Methods and Programs in Biomedicine*, vol. 190, Article ID 105361, 2020.

[18] J. Su, Y. Liu, and J. Wang, "Ultrasound image assisted diagnosis of hydronephrosis based on CNN neural network," *Journal of King Saud University Science*, vol. 32, no. 6, pp. 2682–2687, 2020.

[19] Z. Liu, C. Yang, J. Huang, S. Liu, Y. Zhuo, and X. Lu, "Deep learning framework based on integration of S-Mask R-CNN and Inception-v3 for ultrasound image-aided diagnosis of prostate cancer," *Future Generation Computer Systems*, vol. 114, pp. 358–367, 2021.

[20] R. M. Dewi and U. N. Wisesty, "Classification of polycystic ovary based on ultrasound images using competitive neural network," *Journal of Physics: Conference Series*, vol. 971, no. 1, Article ID 12005, 2018.

[21] B. Cahyono, M. S. Mubarok, and U. N. Wisesty, "An implementation of convolutional neural network on PCO

classification based on ultrasound image," in *Proceedings of the 2017 5th International Conference on Information and Communication Technology (ICoIC7)*, pp. 3–6, Melaka, Malaysia, May 2017.

[22] K. Maheswari, T. Baranidharan, S. Karthik, and T. Sumathi, "Modelling of F3I based feature selection approach for PCOS classification and prediction," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 1, pp. 1349–1362, 2020.

[23] B. Purnama, U. N. Wisesti, F. Nhita, A. Gayatri, and T. Mutiah, "A classification of polycystic Ovary Syndrome based on follicle detection of ultrasound images," in *Proceedings of the 2015 3rd International Conference on Information and Communication Technology (ICoICT)*, pp. 396–401, Nusa Dua, Indonesia, May 2015.

[24] C. Gopalakrishnan and M. Iyapparaja, "Detection of polycystic ovary syndrome from ultrasound images using SIFT descriptors," *Bonfring International Journal of Software Engineering and Soft Computing*, vol. 9, no. 2, pp. 26–30, 2019.

[25] B. Potočnik, J. Munda, M. Reljic et al., "Public database for validation of follicle detection algorithms on 3D ultrasound images of ovaries," *Computer Methods and Programs in Biomedicine*, vol. 196, Article ID 105621, 2020.

[26] S. Bharati, P. Podder, and M. R. Hossain Mondal, "Diagnosis of polycystic ovary syndrome using machine learning algorithms," in *Proceedings of the 2020 IEEE Region 10 Symposium. TENSYMP*, pp. 1486–1489, Dhaka, Bangladesh, June 2020.

[27] M. S. Khan Inan, R. E. Ulfath, F. I. Alam, F. K. Bappee, and R. Hasan, "Improved sampling and feature selection to support extreme gradient boosting for PCOS diagnosis," in *Proceedings of the 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*, pp. 1046–1050, Las Vegas, NV, USA, January 2021.

[28] J. J. Cheng and S. Mahalingaiah, "Data mining polycystic ovary morphology in electronic medical record ultrasound reports," *Fertility Research and Practice*, vol. 5, no. 1, pp. 13–17, 2019.

[29] H. Li, J. Fang, S. Liu et al., "CR-unet: a composite network for ovary and follicle segmentation in ultrasound images," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 4, pp. 974–983, 2020.

[30] C. Carvalho, J. Silva, and A. Campilho, "Deep learning approaches for gynaecological ultrasound image segmentation: a radio-frequency vs. B-mode comparison," in *Proceedings of the Image Analysis and Recognition*, pp. 295–306, Berlin, Germany, June 2019.

[31] N. Singhal, S. Kudavelly, and G. A. Ramaraju, "Deep learning based junctional zone quantification using 3D transvaginal ultrasound in assisted reproduction," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2133–2136, Montreal, Canada, July 2020.

[32] A. G. C. Pacheco and R. A. Krohling, "The impact of patient clinical information on automated skin cancer detection," *Computers in Biology and Medicine*, vol. 116, Article ID 103545, 2020.

[33] G. Holste, S. C. Partridge, H. Rahbar, D. Biswas, C. I. Lee, and A. M. Alessio, "End-to-end learning of fused image and non-image features for improved breast cancer classification from mri," in *Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, pp. 3287–3296, Montreal, Canada, October 2021.

[34] S. E. Spasov, L. Passamonti, A. Duggento, P. Lio, and N. Toschi, "A multi-modal convolutional neural network framework for the prediction of alzheimer's disease," in *Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 1271–1274, Honolulu, HI, USA, July 2018.

[35] K. W. M. K. Wang, "Transfer learning techniques," in *Big Data Technologies And Applications*Springer, Berlin, Germany, 2016.

[36] D. Jia, W. Dong, R. Socher, Li-J. Li, K. Li, and L. Fei-Fei, "Imagenet: a large-scale hierarchical image database," in *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255, Miami, FL, USA, June 2009.

[37] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015*, pp. 1–14, San Diego, CA, USA, May 2015.

[38] Q. Zhang, "Deep learning frameworks for computer aided diagnosis based on medical images," *Dissertation abstracts international. B, The sciences and engineering*, vol. 82, no. 2, 2021.

[39] C. Szegedy, "Going deeper with convolutions," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1–9, Boston, MA, USA, June 2015.

[40] C. Szegedy, V. Vanhoucke, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2818–2826, Las Vegas, NV, USA, June 2016.

[41] N. Dong, L. Zhao, C. H. Wu, and J. F. Chang, "Inception v3 based cervical cell classification combined with artificially extracted features," *Applied Soft Computing*, vol. 93, Article ID 106311, 2020.

[42] A. G. Howard, M. Andreetto, and H. Adam, "Mobilenets: efficient convolutional neural networks for mobile vision applications," 2017, https://arxiv.org/abs/1704.04861.

[43] E. Suharto, A. P. Widodo, and E. A. Sarwoko, "The use of mobilenet v1 for identifying various types of freshwater fish," *Journal of Physics Conference Series*, vol. 1524, pp. 1–6, 2020.

[44] G. Huang and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the 2017 IEEE Conference Computer Visible Pattern Recognition*, Honolulu, HI, USA, June 2016.

[45] A. Shazia, T. Z. Xuan, J. H. Chuah, J. Usman, P. Qian, and K. W. Lai, "A comparative study of multiple neural network for detection of COVID-19 on chest X-ray," *EURASIP Journal on Applied Signal Processing*, vol. 2021, no. 1, pp. 50–16, 2021.

[46] P. Bansal, R. Kumar, and S. Kumar, "Disease detection in apple leaves using deep convolutional neural network," *Agriculture*, vol. 11, no. 7, p. 617, 2021.

[47] E. F. White, "Data fusion lexicon," *Technical Panel C3*, vol. 15, 1991.

[48] F. Castanedo, "A review of data fusion techniques," *The Scientific World Journal*, vol. 2013, Article ID 704504, 19 pages, 2013.

[49] S. C. Huang, A. Pareek, S. Seyyedi, I. Banerjee, and M. P. Lungren, "Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines," *NPJ Digital Medicine*, vol. 3, no. 1, 2020.

[50] C. T. Duong, R. Lebret, and K. Aberer, "Multimodal classification for analysing social media," 2017, https://arxiv.org/abs/1708.02099.

[51] A. Patil and M. Rane, "Convolutional neural networks: an overview and its applications in pattern recognition," *Information and Communication Technology for Intelligent Systems*, vol. 195, pp. 21–30, 2021.

[52] G. Litjens, T. Kooi, B. E. Bejnordi et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60–88, 2012.

[53] D. P. Kingma and J. L. Ba, "Adam: a method for stochastic optimization," in *Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015*, pp. 1–15, San Diego, CA, USA, May 2015.

[54] I. F. Thufailah, Adiwijaya, U. N. Wisesty, and Jondri, "An implementation of Elman neural network for polycystic ovary classification based on ultrasound images," *Journal of Physics: Conference Series*, vol. 971, no. 1, 2018.