NeuroPCOS: Detection of Polycystic Ovary Syndrome in Ultrasound Images Using Filter and Transfer Learning model

MSc Research Project
Data Analytics

Saheli Dutta
Student ID: x21246513

School of Computing
National College of Ireland

Supervisor: Sasirekha Palaniswamy

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Saheli Dutta
Student ID:	x21246513
Programme:	Data Analytics
Year:	2023
Module:	MSc Research Project
Supervisor:	Sasirekha Palaniswamy
Submission Due Date:	14/12/2023
Project Title:	NeuroPCOS: Detection of Polycystic Ovary Syndrome in Ultrasound Images Using Filter and Transfer Learning model
Word Count:	6982
Page Count:	23

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	14th December 2023

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

NeuroPCOS: Detection of Polycystic Ovary Syndrome in Ultrasound Images Using Filter and Transfer Learning model

Saheli Dutta
x21246513
MSc Research Project
National College of Ireland

Abstract

In recent years, polycystic ovary syndrome (PCOS) has been seen a lot in women, because of their lifestyle and food habits. This particular syndrome is affecting their health to a large extent. Keeping these diseases undiagnosed can aggravate the possibility of cancer and damage the reproductive system at any stage of age. There have been a lot of studies conducted related to this topic where several kinds of deep neural networks have been employed to diagnose and identify PCOS-infected ultrasound images because the ultrasound images help to see the condition of the ovary, like a CT scan for the brain or MRI. Unfortunately, because of the capturing devices, the images are exposed to noise, which can lead to the wrong diagnosis of the disease. That is why introducing the filtration method is very necessary. This proposed research utilized the efficacy of a median filter to cancel the noises and fed the images to transfer learning models such as VGG16 and ResNet50 because of their deep hierarchical architecture and skip connection, respectively. Moreover, due to data limitation issues, the ImageDataGenerator data augmentation method is utilized as well. Among these models, ResNet50 outperformed VGG16 by showing balanced specificity(46%) and sensitivity(43%), indicating that identifying both positive and negative classes. Although VGG16's (81%) accuracy is better than the ResNet50 model(52%) VGG16 model fails to identify true positive cases with a low sensitivity of 12% which indicates biases in the model and makes the model unreliable for PCOS diagnosis.

Keywords: Medical Diagnosis, Median Filter, Data Augmentation, Transfer Learning Model.

1 Introduction

In the 21st century, the most pronounced disease for women is Polycystic Ovary syndrome (PCOS), which occurs due to massive hormone imbalances in the body. Mainly, women from 18 to 44 years old are diagnosed with this particular kind of disease during their reproductive years (Morang et al.; 2019). If this disease is left untreated, then it can lead to permanent infertility or ovarian cancer. Normally, ovaries with cysts are bigger in size and contain large amounts of follicles. Excessive hair growth, Periods with irregularities, and problems in pregnancy are early indications of being diagnosed with PCOS [1]. The exact cause behind polycystic ovaries is still unknown, but hormonal imbalances and insulin resistance are the major ones. Resistance to insulin means the body is rejecting the insulin; that's why more insulin will be induced which in turn causes the body to generate more testosterone. That is why normal ovulation gets interrupted and hormonal imbalance initiates.

In a traditional scenario, PCOS can be mainly categorized into two kinds [2].

- **Insulin-resistant**: This occurs due to heightened levels of insulin in the body, which triggers more testosterone.
- **Pill-induced**: Too many birth control pills help increase the luteinizing hormone in the body which affects the ovulation cycle.
- **Inflammatory**: Excessive androgen triggers hormonal imbalance and prevents ovulation. Stress and vitamin D deficiency can actually accelerate this condition.
- **Hidden**: This PCOS occurs due to iodine and zinc deficiency; thyroid disease can also be one of the symptoms.

Diagnosing the disease in its early stages is extremely crucial; therefore, in the medical domain, the utmost importance has been given to ultrasonographic images, which can visually help to identify the affected ovary and healthy ovary within a significant short span of time. This medical image captures cysts in the ovary from various angles. In today's world, artificial intelligence has been utilized a lot to identify and segment images with the help of deep network models, which can further help to make decision to early prediction of this kind of disease in the clinical world.

1.1 Background

Drawing from the Bharathi et al. (2017) study, lifestyle and stress are the biggest contributors to having PCOS, and that is why rural women are less inclined to have it compared to metropolitan women. Therefore, the way of leading life should be changed to avoid such diseases. In another study by the Rotterdam Consensus [3], if there are more than 12 follicles present in the ovaries, then it should be considered polycystic ovaries, and this can be an early sign of PCOS. Statistic provided by WHO [4], 70% women get no proper medication or diagnosis and remain untreated which can gradually led to type 2 diabetes, infertility, and endomterial cancer. Figure 1 Shows the difference between normal and infected ovaries. However, according to the research by (Hassan and Mirza; 2020), to diagnose PCOS with precision and accuracy, artificial intelligence can be used to achieve a prime diagnostic result. Especially for PCOS, ultrasonographic images are very helpful because they can capture enlarged cysts in the ovaries, which are mostly 2–9 mm in size. The authors have employed several machine-learning algorithm to diagnose the particular disease.

Unfortunately, the problem is that those medical images sometimes become very noisy, and this noise can actually omit useful information; hence, denoising medical images and

[1] https://www2.hse.ie/conditions/polycystic-ovary-syndrome/
[2] https://www.indiraivf.com/blog/types-of-pcos
[3] https://www.fertstert.org/article/S0015-0282(03)02853-X/fulltext#secd572228e361
[4] http://surl.li/ntjwi

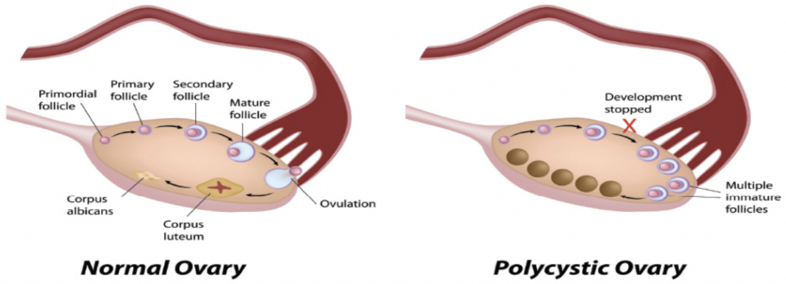

Figure 1: Visual representation of normal ovary and infected ovary (Neuzil; 2014)

then feeding them into algorithms is very necessary to achieve successful prediction. As indicated by (Motwani et al.; 2017), most of the noises are generated by data channels, radiation from various sources, and capturing devices. To denoise the images, the authors have proposed spatial or transform domain filtering methods based on the type of noise, Spatial filters fix the bandwidth at a very low-frequency level, which is especially based on high-frequency noise. The transform domain filters are a combination of adaptive and non-adaptive filters where the Fourier transform and complex mathematical calculations are used. The drawback of transform domain filters is that they are very time-consuming and complex compared to linear and non-linear filter processes. Figure 2 shows types of filters.

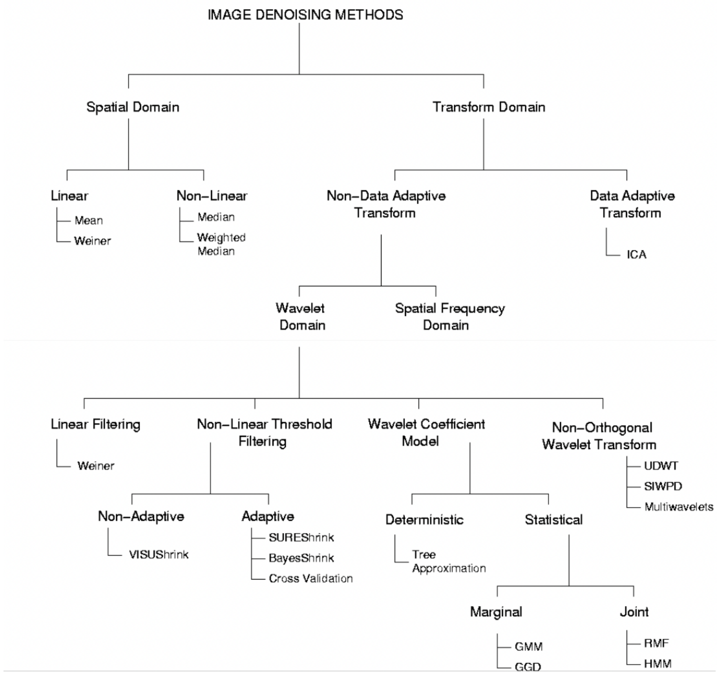

Figure 2: Image De-Noising Methods (Motwani et al.; 2017)

In their study, (Jasim et al.; 2019) found that there are four types of noises which are

Gaussian, Salt & Pepper, Spcekle, and Poison noise. Mainly, Gaussian and Spcekle noise are commonly found in electronic images and medical images as well. Speckle noise mostly blurs the edges and fine details in ultrasound and CT images. Hence, to capture proper, meaningful information from medical images and to prevent loss of information, spatial domain filtering techniques are necessary. Now, coming to spatial domain filters, the mean filter works very fast but damages the sharp edges and loses information on them. On the other hand, the median filter can remove the noise very reasonably by employing low-pass filters on high-frequency pixel groups. Hence, the removal of distorted pixels from the ultrasonographic images is the prime focus of this proposed study.

Noise-free images can be fed into transfer learning models to classify infected and non-infected ovaries because these models are trained on large datasets with 1000 different image classification tasks [5]. The deep architecture is crucially helpful to capture important patterns and features from medical images with precision and accuracy. For instance, the InceptionV3 model is employed to diagnose malignant tissue in the breast and has exceptional outcomes (Chang et al.; 2017).

1.2 Motivation

To this day, diagnosing and detecting cysts in their early stages is still challenging. Instead of having several studies on this particular topic, a proper diagnosis method has not yet been discovered. Also, proper imaging in ultrasonographic photos is crucial to detecting PCOS. If the image-capturing instrument produces noise along with capturing the photo, then the final result will lead to misinterpretation. This proposed thesis will focus on the use of a non-linear(median) filter to produce a better noise-free image; furthermore, with the help of the pre-trained model, those medical images can be diagnosed more efficiently. The primary reason behind utilizing median filter to suppress noise and maintain information around edges with minimal complexity is to highlight the salient features that are important for decision-making, which can help deep learning model to achieve good amount of accuracy while diagnosing ovarie's ultrasonographic images.

As indicated by Choubey et al. (2021), to extract the pattern of noise, the researcher utilized a discrete wavelet transform (DWT) filter in it's primary phase. On top of that, median filters were used in the second step of research to gain an understanding of noise forms, and with the help of a deep learning model, they achieved successful detection of polycystic ovaries from ultrasound images. Also, (Hijab et al.; 2019) demonstrated how well the Visual Geometry Group (VGG16) and customised CNN model can diagnose malignant tissue in breasts and achieved impressive results. Transfer learning models such as VGG16, Inception V3, and Residual Network (Resnet50) have already been used so many times in the medical domain to diagnose disease from medical images due to their shorter training times. All of these factors became the driving force behind the presented thesis where the main focus would be-

- A median filter will be employed to reduce distorted noise from ultrasonographic images because a median filter can retain useful information without blurring too many of the sharp edges, unlike a mean filter. Also, median filtering is very easy to implement, which means it's not mathematically complex and less time-consuming, unlike transform domain filtering techniques.

[5] http://surl.li/jnrhc

- VGG16 and ResNet50, both of these transfer learning model will be employed to compare the accuracy and precision between them. These both models have sufficient model complexity to handle the fine-detailing of medical images and also they have powerful feature extraction ability to learn hierarchial features in abstract level [6], [7].

1.3 Research Goal

The major objective of this study to add median filtering method to the medical images to get better visibility of pattern which can enhance the accuracy of transfer learning models and this can lead to better diagnosis of chronic disease like PCOS. The flow of this proposed research are the following -

1. Implement median filter to remove distorted pixels from the image.
2. Augment the filtered data with the help of Keras Image data generator to create more sample during training because it will help to generalize the model.
3. Utilize transfer learning model with deep architecture such as VGG16 and ResNet50 to those filtered, augmented ultrasonographic images and compare the result between them.
4. Try to identify which model performed better by predicting infected or non-infected ovaries on unseen images.

1.4 Research Question

- How well does the ResNet50 model perform compared to the conventional VGG16 model to detect PCOS images with the help of a median filter?

1.5 Report Structure

The structure of this study will be as follows - Section (2)provides a review of the corresponding literature and helps to conclude the contribution of those studies toward the medical domain. Coming to section (3), the methodology of this study is discussed in detail which starts from business understanding and data collection to data pre-processing and model building. Section (4), provides more details on the workflow, median filter, and transfer learning model specification. Section (5) covers the implementation of the methodology, which starts from the data load and filter and pre-trained model implementation to model deployment. In section (6), results are discussed with proper evaluation metrics to understand how deep learning models and filters work. Lastly, Section (7) concludes the whole research, and future scopes are discussed here.

2 Literature Review

In recent years, there has been a lot of research in the medical domain where some researchers applied traditional neural networks to extract and identify the salient features

[6]`https://keras.io/api/applications/vgg/`
[7]`https://keras.io/api/applications/resnet/`

to diagnose any disease, whereas other scholars implemented image de-noising methods to capture useful pixels in the medical data and employed that in deep learning models. Based on these, some papers are discussed in this section which influenced this proposed thesis. The entire literature review is partitioned into three subcategories (2.1 and 2.2).

2.1 Types of Filtering Method in Image Processing Domain

In accordance with the finding of Jasim et al. (2019), mainly Mean, Median, and Weiner filters are utilized to identify the type of noises. With the help of evaluation metrics, comparisons are also drawn to understand which filters perform well. Based on the researcher's findings, Speckle noise is commonly found in medical images such as CT scans, MRI, and Ultrasonographic images. In terms of filters, the median filter calculates the closest pixel value which is close to the damaged pixel based on the median method, and substitute the damage pixel value with the calculated one. But the mean filter calculates the average which smokes the sharp edges and loses information too much. Gaussian, speckle, salt & pepper, and poison noises are imposed on the image in three different ratios. While comparing the performances, the median filter works the best to identify Gaussian noise and achieves a 26.03 PSNR value. It is clear the median filter can be used to remove noise in medical images because this filter can help preserve fine detailing while removing damaged pixels. To diagnose PCOS, this proposed thesis will utilize a median filter to experiment with how this can help enhance the performance of deep learning models.

Like the previous research, Vijayan and Venkatachalam (2021) also conducted a similar kind of experiment where the authors used Discrete Wavelet Transform and KSVD filtering methods. These filters follow a completely different approach while removing the noise, KSVD is more of like an iterative method that uses a sparse process to process the image. On the other hand, DWT filters use both high and low pass filters to break the image pixel by pixel to give the image a multi-abstract structure. Unfortunately, both of these filters have major flaws such as complex processes while correcting the dictionary, which affects the efficiency while processing large images and very slow convergence. However, the KSVD filter works much better than DWT based on the evaluation metrics. This proposed thesis will avoid this complex filtering method because the aim is to speed up the cyst-detecting process in ultrasonographic images so employing these kinds of filters will slow the diagnosis.

In another research proposed by Brindha and Rajalaxmi (2023), authors followed the traditional method of using a Non-Local means filter to remove noise from the ultrasound images. NLM filter subdues the noise using the weighted average value of distorted pixels. Afterward, images are fed into eight different machine-learning models to compare the accuracy. However, Choubey et al. (2021) detects PCOS from images by using filtering methods in two different stages. Firstly, the scholars tried to understand the noise pattern using DWT filters, and then based on the outcome of DWT, noise patterns are recognized. In the second stage median and winter filters are employed to denoise the medical images. After filtering the images twice, Mean Square Error (MSE) came down to 97% which is impressive because this can help to diagnose noisy images.

Wieclawek and Pietka (2019) demonstrated the utilization of granular computing with crisp and fuzzy filtering variants. These filtering techniques are employed to reduce the noise in 2496 US breast images and 15 CT images. The crisp filtering method utilizes binary information for morphological reconstruction, Conversely, fuzzy filtering techniques use fuzzy membership functions to include pixels. The Crisp granular filter removes nominal noise and the fuzzy granular filter extracts strong noise without clouding boundaries. These approaches are compared with traditional filters like mean, median, and Weiner. In the evaluation process, it is seen that these complex filtering methods are working much better than traditional filters but they are computationally very expensive and more time-consuming.

This proposed research aims to balance the proportion between lesser execution time and suppress noises from medical images to such an extent that salient features can be easily noticed.

2.2 Transfer Learning Models Utilised in Medical Field

Based on the research by Abbas et al. (2020), the Medical image classification task is advanced by using deep learning models, and in this research, they followed the class decomposition approach. Mamograms, Colorectal Cancer, and Chest X-ray images are utilized to evaluate their model. The proposed DeTraC model accomplishes 99.80% accuracy while pre-trained models like VGG-16, and ResNet achieved 98.5% on those medical images. According to the research, the more the number of subclass, the better convergence time and more accuracy. This research proves the usefulness of the pre-trained model's robustness in the medical domain to achieve better results in a limited time.

Similarly, In their study Hijab et al. (2019), scholars tried to diagnose breast cancer from ultrasound images using a deep learning model. They have employed different kinds of models, The first is the custom CNN model. Secondly, the VGG16 transfer learning model where hyperparameter tuning is done to get rid of overfitting issues and then train the model on those medical images. In this VGG16 model, hyperparameter tuning is done by applying stochastic gradient descent(SGD) to update weights in backpropagation. Training of VGG16 ran for 50 epochs and 1000 medical images and also to generalize the model and overcome the class imbalance problem data augmentation (flipping, rescaling, zooming, etc) are done. During evaluation, VGG16 reached up to 97% accuracy while the custom CNN model didn't perform well. In recent years, more priority shift towards utilizing transfer learning models because these models have proven it's capacity to reach more accuracy compared to the traditional model, based on that Hosain et al. (2022) also utilized InceptionV3 and custom CNN model and fine tune both of them to get optimum result. Moreover, the researcher here employed Kera's ImageDataGenerator to augment the images to overcome data imbalance and limitation problem. The InceptionV3 model is adjusted by extracting the top most layers and putting two additional layers to detect binary classes. On the contrary, the CNN model had 3x3 filters for each convolutional layer and one sigmoid activation function in the final layer while Rectified Linear Unit (ReLU) activation function used in the hidden layer. During evaluation, Both the custom CNN model and Inception V3 model achieve optimum accuracy (98.12% and 96.56% respectively).

Another remarkable research was conducted by Brindha and Rajalaxmi (2023), where VGG16, ResNet50, and CNN models are utilized to diagnose polycystic ovaries. In this research, wights of transfer learning models are used in ANN and Support Vector Machine (SVM) classifiers and the last layers of pre-trained models are exchanged by classifiers. Along with that, the original architecture of VGG16 and ResNet50 are also utilized. During comparison between these models, VGG16 outperformed by reaching accuracy up to 93%, where based on the F1 score custom CNN model and VGG16 model both achieved almost similar results. Again, the same kind of research has been conducted by Chang et al. (2017) to diagnose breast cancer. Here, data augmentation techniques such as flipping, shifting, and rotating are done to get rid of data imbalance issues. On top of that, the Inception V3 model is utilized to reduce training time and faster convergence. During the evaluation of the model, the accuracy reached up to 89% for the infected class and 83% for non-infected.

It is now proven that the utilization of the transfer learning model, especially in the medical domain has increased a lot in recent years. That's why this proposed study aims to combine the effectiveness of the filter and transfer learning models' deep architecture to achieve faster diagnosis with better accuracy.

3 Methodology

This proposed research utilizes the de-noising method using a median filter and two transfer learning models (VGG16 and ResNet50) to classify and detect cystic ovaries from medical images. Removing deformed pixels from images and keeping the salient features highlighted in ultrasound images is the major task to help the transfer learning models achieve their accuracy. Based on the evaluation results, a comparison is to be drawn on the performance of both models. In-depth architecture and transfer learning knowledge of these pre-trained models help to boost the training to a great extent, in the medical domain these models have already shown noteworthy potential in improving PCOS detection. The KDD (Knowledge Discovery in Databases) process is heeded in this research as it follows every step that is essential to perform this research.

3.1 Business Understanding

In accordance with Yang et al. (2015), Knowledge Discovery in Databases (KDD), this particular approach has been utilized a lot in recent years to gain valuable insight by mining data.

Figure 3 shows the steps involved in KDD methodology [8]. To categorize and extract features from data to identify valuable structured patterns that lie under the raw form of the data, KDD is an ideal process which is why it is already employed in many advanced machine learning projects. KDD consists of several crucial stages like domain knowledge, data collection, pre-processing of data, data transformation, data mining, visual evaluation, and interpretation of the data. The main advantage of this approach is data mining, which integrates databases, statistics, neural networks, etc. Hence, the KDD approach is followed in this research as a powerful tool for decision-making and demand analysis.

3.2 Data Collection

The entire study starts by collecting ultrasound images from the Kaggle dataset [9]. The repository has separate train and test folders where 2288 healthy and 1568 infected ovaries

[8] https://miro.com/app/board/uXjVPgwFBZs=/

[9] https://www.kaggle.com/datasets/anaghachoudhari/pcos-detection-using-ultrasound-images

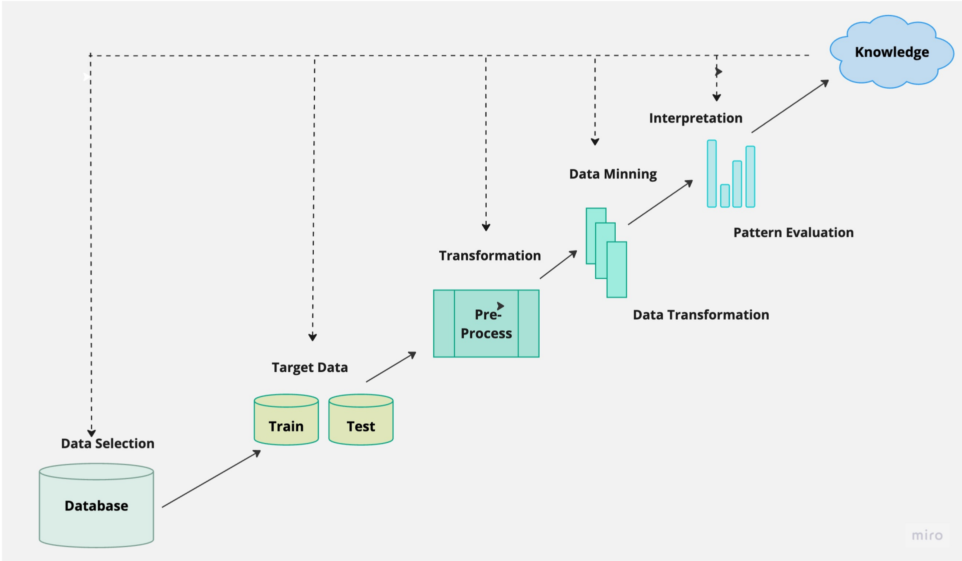

Figure 3: KDD Diagram

are present. Under the train directory, 781 ultrasound images are infected, and 1143 images are non-infected.

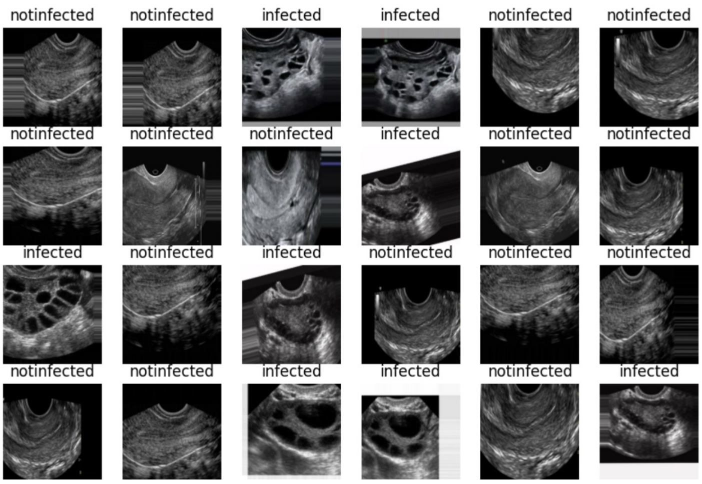

Figure 4: PCOS and Non-PCOS Data

Similarly, in the test folder, the number of healthy ovaries is 1145, and unhealthy ovaries are 787. The infected ovaries are enlarged in shape with a growing number of follicles, while the non-infected ovaries are in their original shape with no abnormalities. Figure 4 represents the visuals of infected and non-infected ovaries from the Kaggle repository.

3.3 Data Pre-Processing

During this stage, the most crucial step is to remove noise and retain all-important features in the ultrasound images. To achieve that Median filter is employed. As indicated by Omer et al. (2018), the median filter successfully suppressed salt & pepper noise from the CT scan images while edges are preserved with sufficient noise immunity. However, while conducting the literature review, the Mean filter is also used for the same purpose but the edges are blurred too much and that became the reason for losing important information. Hence, this proposed thesis focuses on using a Median filter on ultrasound images to highlight the important feature by removing distorted pixels with an easy implementation process. The OpenCV library is employed to utilize the median filter as it provides robust and efficient performance [10]. Moreover, a NumPy function is also created to use the filter through Open CV and medianBlur functions. Figure 5 illustrates the visuals of filtered images.

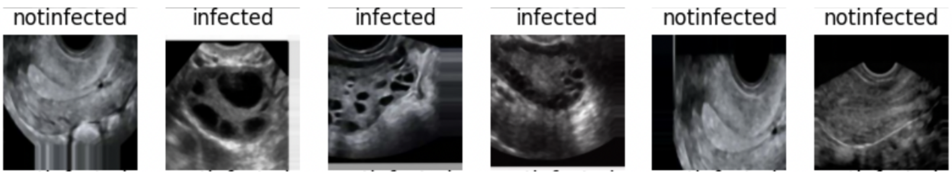

Figure 5: Filtered Data

Apart from the filtering method, data normalization and data binarization are also done to boost efficiency. Initially, a Tensorflow dataset is created to classify the images where binary classes are in the form of 0 and 1. The classes are infected and not infected respectively. Data normalization is done by dividing the pixel value by 255, because pixels in the image range from 0 to 255, to normalize the pixel value which ranges from 0 to 1, a pixel value divided by 255 is needed to ensure all the features are in the same scale, also easy to fed into deep learning models as well.

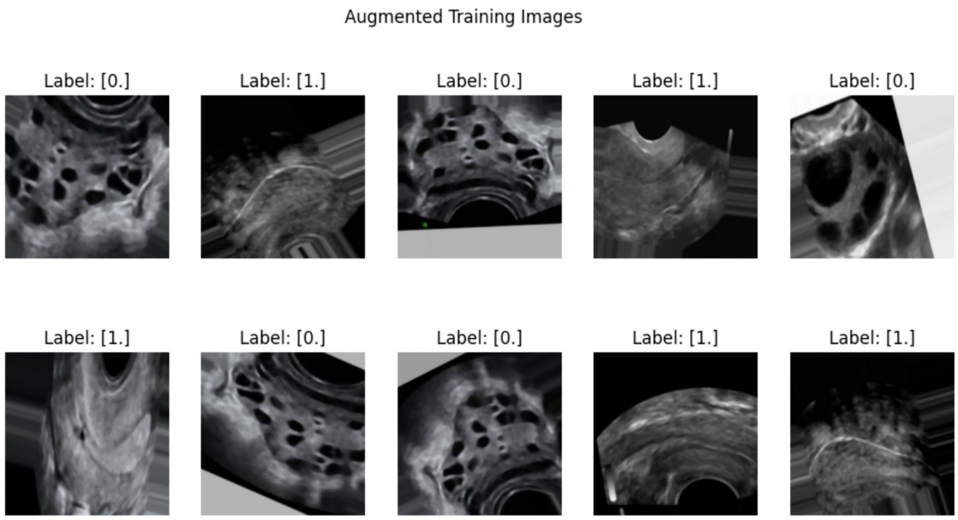

Figure 6: Augmented Training Data

[10]https://docs.opencv.org/3.4/dc/dd3/tutorial_gausian_median_blur_bilateral_filter.html

Lastly, data augmentation is done to broaden the diversity of the training dataset to get rid of the class imbalance problem and improve the generalization capacity of the transfer learning model. To augment the data, the ImageDataGenerator class is utilized from the TensorFlow Keras library and the reason behind using this particular approach is to generate sets of ultrasonographic image data with real-time data augmentation [11]. Figure 6 displays the augmented training images.

This technique is especially promising for several causes. Rather than keeping augmented images stored, this real-time data augmentation is applied on the go during training, saving a significant amount of storage space. Data augmentation is done by zooming, shifting, rotating, and flipping to make the model more robust. Data flow generators are employed to supply a constant stream of augmented sets of images during the training stages. During the validation and testing phase, only image rescaling is done to avoid any kind of synthetic variations.

3.4 Modeling

The main purpose of this presented study is to build a robust diagnostic model where diagnosis of PCOS can be done effectively, even in the presence of noisy medical images. This research aims to overcome the issues that generally happen due to the presence of noise in the images and also to assure accuracy and a dedicated PCOS diagnostic process. There are lots of studies that have utilized this notional deep learning framework, this study puts forward an additional step which is the de-noising image phase, to ensure the transfer learning model can be utilized and produce accuracy without dealing with additional noise in the image.

Here, the median filter is used in the data pre-processing step to enhance the pixel-to-pixel quality of the ultrasound images. Reducing the noise while preserving the information around sharp edges and not blurring too much feature information are the main advantageous points of employing the median filter through OpenCV. Also, with that, all the images are rescaled to 0 and 1 so that all the feature comes into the same scale, and labeling both of the classes to 0 and 1 are done to pre-process the data, all of these help to feed the images into transfer learning models. Figure 7 illustrates the workflow of the proposed research.

Data augmentation is done using ImageDataGenerator to generalize the model to unseen data and get rid of data imbalance problems as well, VGG16 and Resnet50 both of these transfer learning models are employed to classify the PCOS and NON-PCOS images. Lastly, their performance is evaluated through evaluation metrics such as accuracy, precision, recall, specificity, and sensitivity, also, prediction has been made on the test dataset to see how both of the models predict the classes.

4 Design Specification

In the design specification, there are two sub-sections, which cover the key aspects of this research. This section will focus on the median filter and transfer learning models (VGG16 and ResNet50) specifications.

[11] `https://www.tensorflow.org/api_docs/python/tf/keras/preprocessing/image/ImageDataGenerator`

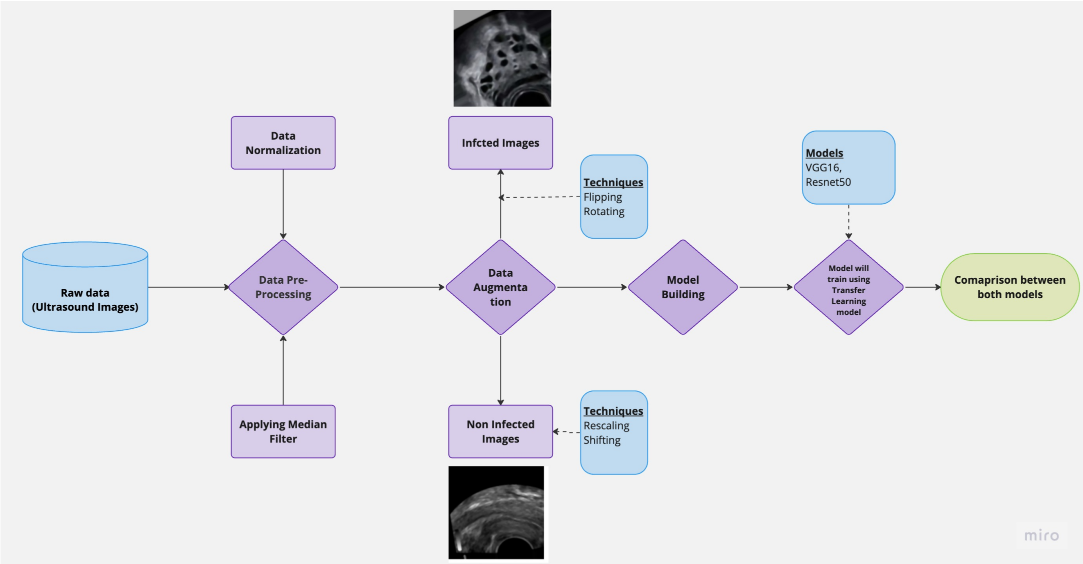

Figure 7: Methodological Workflow

4.1 Median Filter Specification

The median filter is a commonly used method to remove speckles and Gaussian noise from images. Mainly, in the medical domain noises in images come from capturing devices where some pixels are set to a minimum value, on the other hand, some of the pixels are set to maximum, creating noise in the image (Goel et al.; 2019). The median filter processes the neighbor pixels that surround the noisy or damaged pixels also known as the Kernel area, this filter calculates the median value of the kernel area and replaces the distorted pixel's value with that. The amount of recovery from noise can be measured by the Peak signal-to-noise ratio (PSNR) value, Mean Square Error (MSE), and Root Mean Square Error (RMSE). Figure 8 explains the workflow of the median filter algorithm.

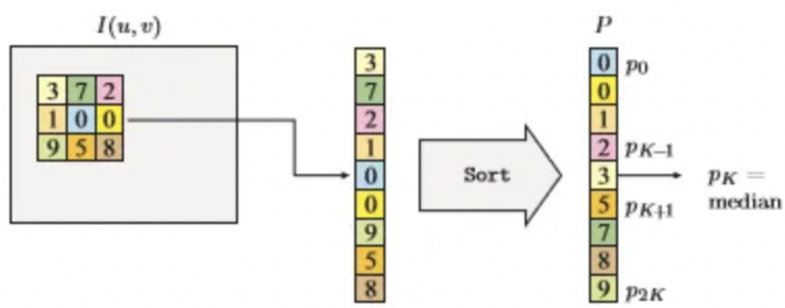

Figure 8: Median Filter Working Diagram (Nattadet; 2018)

4.2 Transfer Learning Model Specification

Two transfer learning models are employed in this study. Both these models have utilized knowledge from a pre-trained network, which means these models are already trained on

1000 images from the imagenet database to classify 1000 different images.

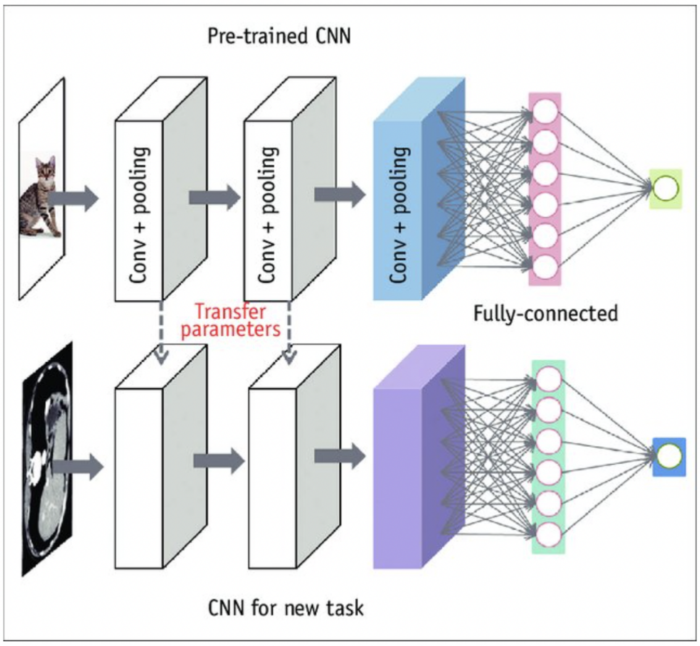

Figure 9: Tranfer Learning Model Flow Diagram (Do et al.; 2020)

This proposed research utilized those models and made changes in the output layer from 1000 neurons to 2 neurons as the task is binary [12]. Figure 9 explains the methodology behind the transfer learning model. The reasons behind employing these kinds of models are listed below -

- Transfer learning model is already trained on a large and generalized dataset, this helps the model to be robust enough to classify all the image types.
- Transfer learning can quicken the training process because of their comprehended hierarchical features from the large dataset. As the model is already built with pre-trained weights, a significant amount of time is saved compared to building from scratch. This actually contributes to process optimization.

4.2.1 ResNet50 Model Specification

The residual Network is one of the most commonly utilized deep neural network models which is based on a pre-trained network, this particular model is employed a lot in image classification tasks. This CNN-based model has 50 stacked layers on top of each other. The advantageous part of this model is that it deals with the vanishing gradient issue because vanishing gradient was a hard obstacle while training a deep neural network. This deep architecture also has short connections, known as ”skip or Residual connections”, this means the model can skip some connections to learn more deep architectural features and put more emphasis on essential execution to have more effective [13]. Figure 10 shows the deep architecture of ResNet50 with maxpool, convolutional layer, identity block, ReLU activation function, fully connected layer, etc.

[12]https://www.tensorflow.org/tutorials/images/transfer_learning
[13]https://medium.com/@nitishkundu1993/exploring-resnet50-an-in-depth-look-at-the-model-architectur

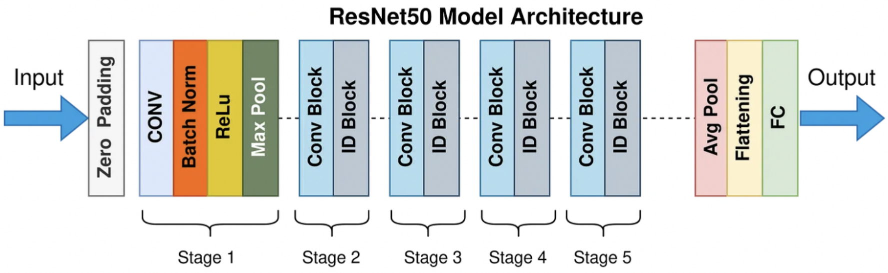

Figure 10: ResNet50 Architecture

4.2.2 VGG16 Model Specification

Visual Geometry Group (VGG) is a variety of convolutinal neural networks, one of the best models for computer vision tasks. This model was invented in 2014 by Andrew Zisserman and Karen Simonyan from Oxford University in imagenet challenge, where the researchers accomplished to identify 1000 different images with 1000 different classes with an accuracy of 92.7%. VGG16 model consists of 16 layers with weight, which simply means these 16 layers have parameters that are learnable. This model takes input in the form of 224x244 pixels and 3 RGB channels and passes through 13 convolutional layers where 3x3 filters with stride 1 are used. CNN layer and max pooling layer are stacked on top of each other and the number of filters is changed from 64 to 512 from convolutional layer 1 to layer 5 [14]. Figure 11 illustrates the deep architecture of VGG16.

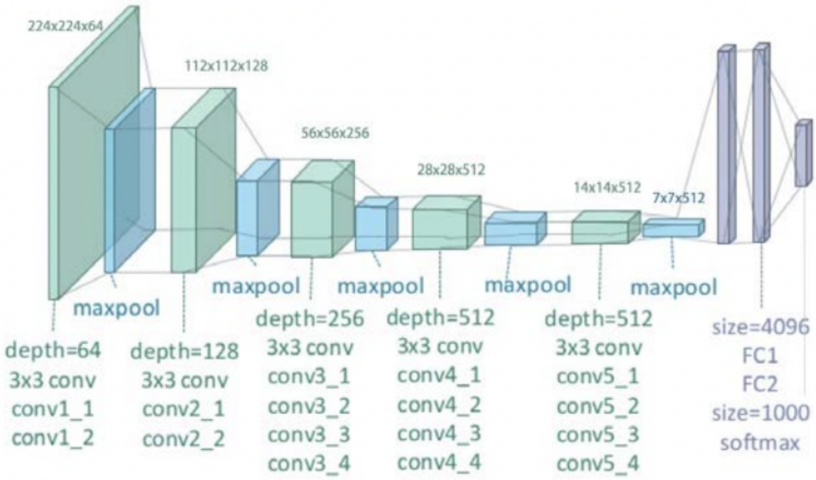

Figure 11: VGG16 Architecture (Zheng et al.; 2018)

5 Implementation

This section is focused on how well the proposed workflow is implemented to achieve the goal. This will start with setting up the machine to carry out the research to load the

[14] https://medium.com/@mygreatlearning/everything-you-need-to-know-about-vgg16-7315defb5918

data and end with filter and deep learning model implementation.

5.1 Enviroment SetUp

The system needs to be set properly to execute the workflow mentioned above. This entire research is done in MacBook Air (M1, 2020 edition) with 8GB RAM, where TensorFlow and Keras library are utilized. The dataset is mounted in Google Colab to pre-process the data further and build the models. Here, transfer learning models are utilized, as a result, Google Colab Pro is used to extend the RAM power and computing units so that seamless study can be carried out. As Keras is a part of the TensorFlow ecosystem, deep neural network APIs are provided which further help to build, train, and experiment with deep architecture models. Hence, ResNet50, VGG16, ImageDataGenerator, and callbacks, all of them are utilized from the same. Along with that to visualize the data and evaluate the model performance the data Matplot, and scikit-learn library are used respectively.

5.2 Data Load

The ultrasound images from the Kaggle repository are downloaded into the local machine and then mounted in the Google Colab to start processing the data. After filtering the data using the median filter, the entire training dataset is divided into training, validation, and testing purposes. Figure 12 shows the number of ultrasound images present in each folder before augmentation. Training and testing datasets are split into 70-30 ratio, out of 30% of the data it was further split to 50% to have validation and test dataset.

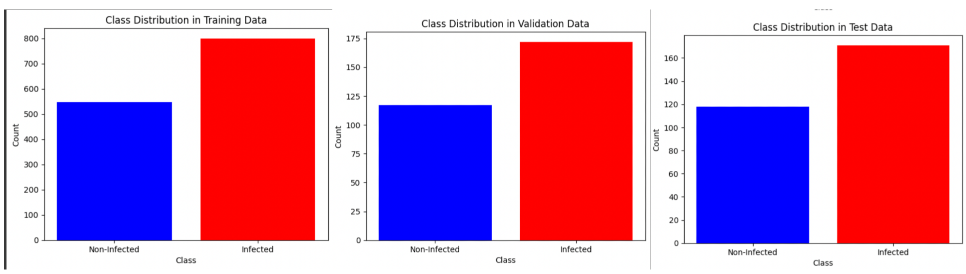

Figure 12: Visual Breakdown of Training, Testing, Validation Datasets

Clearly, the number of images is less, that's why augmentation is applied to the training dataset by zooming, flipping, shifting, and rescaling with the help of ImageDataGenerator from Keras. The flow () method is utilized to generate augmented data in batches during the training process. For validation and testing datasets, only rescaling has been done to scale down all the features on the same scale. Also, this image augmentation process helps to augment data on a real-time basis, which further helps to save up space and optimize the process.

5.3 Filter Implementation

To utilize the filtration method, the proposed architecture incorporates the median filter in the pre-processing stage to de-noise the input images. To achieve the utility of the median filter, a custom TensorFlow function (apply_median_filter_numpy) is applied here with the help of the OpenCV library to get the medianblur function. This particular function confirms the odd size of the Kernel to achieve a filtering effect with proper balance. Lastly, a wrapper is utilized with the help of TensorFlow, this is very crucial to have seamless integration and compatibility, also, a numpy-based filter is required to be included because the medical images are fed into the Deep neural network model.

5.4 Transfer Learning Model Implementation

The proposed study implemented two transfer learning models (VGG16, ResNet50). For ResNet50, the architecture contains ResNet-based networks to segment the binary classes. The model uses pre pre-trained ResNet50 model where the weights are already defined in the Imagenet. Top layers are set to false to use this as a feature extractor, where salient features from the images can be captured. Also, the input shape is fixed to 224 × 224 pixels with RGB channels(3). In particular, global average pooling layers is employed with the ReLU activation function and L2 regularization to extract the important feature from images while preventing any chance of an overfitting problem. Dropout layers are set to 0.3 value, Dense architecture with 2 units, and SoftMax activation function is applied to have binary classification. Figure 13 illustrates the implementation of the ResNet50 model in this research.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 resnet50 (Functional)       (None, 7, 7, 2048)        23587712

 global_average_pooling2d (  (None, 2048)              0
 GlobalAveragePooling2D)

 dense (Dense)               (None, 64)                131136

 dropout (Dropout)           (None, 64)                0

 dense_1 (Dense)             (None, 32)                2080

 dropout_1 (Dropout)         (None, 32)                0

 dense_2 (Dense)             (None, 2)                 66

=================================================================
Total params: 23720994 (90.49 MB)
Trainable params: 133282 (520.63 KB)
Non-trainable params: 23587712 (89.98 MB)
_________________________________________________________________
```

Figure 13: ResNet50 Model Implementation

The compilation process employed an Adam optimizer and binary cross-entropy as the loss function. The model is compiled using the Adam optimizer and binary cross-entropy loss. The model is trained for 200 epochs with EarlyStopping and ModelCheckpoint Callbacks where ModelCheckpoint callback is used to get the best weights based on validation accuracy and EarlyStopping is implemented to stop the training process after 7 epochs if the validation loss doesn't improve. After tuning the parameters several times, this particular model stands out to be the best to identify binary classes.

Coming to the next model, the model architecture of VGG16 consists of convolutional layers to extract features. The pre-trained network of VGG16 is utilized as the base architecture. Similar approach like RedNet50 model is implemented here, weights of the network from Imagenet database is employed and top layers are set to false to identify hierarchial feature from the image. Here, the input shape is set to 224×224 pixels with 3 channels. Following dense layers are used for classification task. The model compilation is done using Adam as optimizer, Binary cross entropy as loss, accuracy as evaluation metrics. Figure 14 shows the implementation of the VGG16 model in this research.

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 vgg16 (Functional)          (None, 7, 7, 512)         14714688

 flatten (Flatten)           (None, 25088)             0

 dense_3 (Dense)             (None, 4096)              102764544

 dropout_2 (Dropout)         (None, 4096)              0

 dense_4 (Dense)             (None, 4096)              16781312

 dropout_3 (Dropout)         (None, 4096)              0

 dense_5 (Dense)             (None, 2)                 8194

=================================================================
Total params: 134268738 (512.19 MB)
Trainable params: 119554050 (456.06 MB)
Non-trainable params: 14714688 (56.13 MB)
_________________________________________________________________
```

Figure 14: VGG16 Model Implementation

Lastly, the model is compiled in the same way as the ResNet50 model. The deep architecture of this model are developed to achieve a harmony between intricacy and generalization, which seeking to achieve precise and robust PCOS detection on ultrasound images.

5.5 Model Deployment

During this phase, the trained models (”lighter_pcos_model”, ”build_regularized_resnet_model”) are saved. Again loaded using the Keras module to test the model’s prediction on the unseen data. The outputs come in the form of probabilities, where higher confidence is given to higher probabilities. This implementation is useful for real-world scenarios displaying its capability to make correct identification on medical images of each patient, which is crucial for functional performance in PCOS detection where ultrasonographic images are fed from outside and the model can predict the class within seconds.

6 Evaluation

This section presents a complete breakdown of the performances of the median filter and transfer learning models. The performance of the median filter is evaluated using Mean Square Error (MSE), Root Mean Square Error (RMSE), and Signal-to-Peak Ratio (PSNR). On the other hand, Accuracy, Specificity, Sensitivity, Recall, and F-1 Score are used to evaluate the performance of Deep neural network models.

6.1 Case Study 1: Use of Filtration on Medical Image

The median filter is implemented in the data pre-processing phase to blur out the noise and make the feature of ultrasound images clear so that the transfer learning model can extract features easily and identify the underlying patterns to identify binary classes. Figure 15 shows the evaluation of the median filter based on the metrics. The outcomes show the respective PSNR, RMSE, and MSE values. The lower PSNR value indicates the quality of the filtered image is decreased a bit while RMSE and MSE with lower values show minimal pixel-based contrast between the original and filtered images. This means the applied median filter successfully reduced the noise from the ultrasound images by trading off the image quality.

```
Average PSNR for Training Set: 11.517660935740853
Average RMSE for Training Set: 0.2791634698030233
Average MSE for Training Set: 0.08577260400190463
```

Figure 15: Median Filter's Performance Based on Evaluation Metrics

Reducing noise and maintaining the important features in the images need to be balanced to optimize the performance, in this research image quality is compromised because of a lower PSNR ratio but the reduction of noise is done successfully by achieving lower RMSE and MSE values.

6.2 Case Study 2: Performance of Pre-trained model on PCOS image classification

6.2.1 Accuracy and Loss Plots

The performance of pre-trained models is visually shown in the graph with the help of the loss and accuracy graph in the training and validation dataset. The graphical presentation gives a complete insight into the model's performance which helps to understand the generalization ability of the model to unseen data. Figure 16 displays the loss and accuracy plot on training and validation data for the ResNet50 model.

The loss graph represents the model's convergence while the accuracy graph indicates the model's ability to identify classes correctly. In the above graphs, within 60 epochs the performance reached to peak then loss started to increase, and because of early stopping callback, the training stopped. The training and validation loss starts from 0.7 and within 60 epochs comes down to values near 0.1 and 0.2. The accuracy graph shows improvement

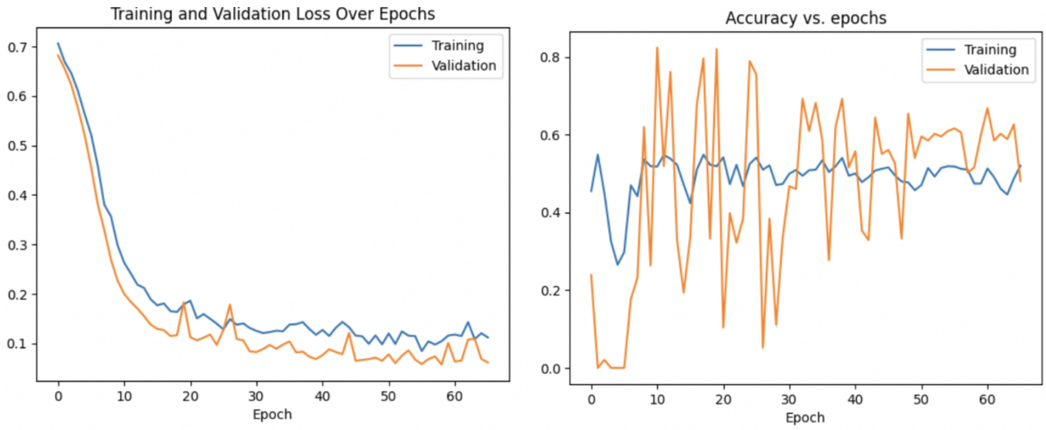

Figure 16: ResNet50 Loss & Accuracy Plot

in the validation dataset from 0.2 to 0.8, while the training set's accuracy goes up from 0.3 to 0.6.

Similarly, the VGG16 model's loss and accuracy plot illustrates the performance of training and validation sets over 30 epochs. Figure 17 displays the loss and accuracy plot on training and validation data for the VGG16 model.

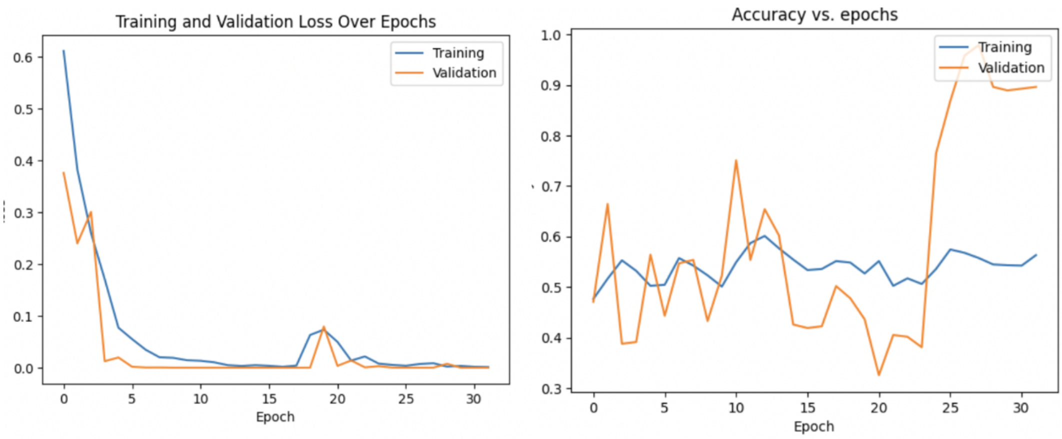

Figure 17: VGG16 Loss & Accuracy Plot

Loss during training comes down from 0.6 to almost 0 within 30 epochs because of early stop callback. In the same way, the validation dataset loss comes to 0 as well over 30 epochs. Accuracy plots show the training set's performance in terms of accuracy increased from 0.5 to 0.6 while the validation dataset's accuracy almost reached 1.0, which indicates a clear overfitting of the model.

Based on the plotted visuals, Resnet50 surpassed VGG16 because of exhibited fast convergence and performed more downward training and validation losses within its respective epoch range. Moreover, the ResNet50 model achieved enhanced accuracy on both datasets compared to the VGG16 model. This establishes the ability to capture more complex features and the generalization power of the ResNet50 model.

6.2.2 Evaluation Metrics

This research is based on the medical domain, where accuracy is one of the crucial factors. VGG16 model achieves an accuracy of 81%, clearly surpassing the ResNet50 model's accuracy of 53%. Nonetheless, a complete analysis of sensitivity and specificity is very critical in this domain and helps to provide a more subtle standpoint. ResNet50 displays a higher sensitivity/recall of 43%, indicating more usefulness in identifying positive cases correctly. Further, the specificity of the ResNet50 model is 46%, indicating a balanced performance in identifying negative instances. ResNet50 model displays its ability to handle both of the classes. Figure 18 shows the evaluation of the model based on the metrics.

Model	Accuracy	Loss	Precision	Recall / Sensitivity	F1 Score	Specificity
ResNet50	0.52	0.03	0.54	0.43	0.47	0.46
VGG16	0.81	0.00	0.58	0.12	0.19	0.88

Figure 18: Model's Performance Based on Evaluation Metrics

On the contrary, the VGG16 model achieves very low sensitivity(12%), which indicates the model's poor performance in identifying true positive cases. This is happening due to overfitting of the model. Moreover, high specificity points to a few false positive cases.

Assuming to diagnosis of any medical problem is very complex, the power of identifying positive classes (PCOS patients) always carries more weight. ResNet50 model is the preferred choice due to its balanced specificity and sensitivity. On the other hand, VGG16 shows higher accuracy, but the balance between specificity and sensitivity is not present. That's why the ResNet50 model is the better choice for accurate, overarching, and robust PCOS diagnosis where both true positive and true negative instances are given an equal amount of focus.

6.3 Discussion

This proposed research was very challenging from the start, as this study deals with medical data which is very hard to arrange to experiment. Secondly, implementing the median filter is the next obstacle because smoothening the edges of the images can cause potential information loss, which can affect the efficacy of the deep learning model while training. With OpenCV and medianblur function, all the ultrasound images are properly filtered without exposing the images too much. Lastly, ResNet50 and VGG16 models are deployed. While evaluating both of the model's performances, ResNet50 displayed a balanced performance compared to VGG16. To deploy the model, both of the pre-trained models are saved and loaded for real-world application, where an infected image is given to both of the models to predict the class. ResNet50 successfully predicts the class and the prediction is interpreted by a dictionary where the probabilities of being "infected" and "notinfected" classes are presented. This exhibits the model's ability to make correct predictions on unseen data, a key factor in its applicability in the real world for PCOS detection.

7 Conclusion and Future Work

Diagnosing PCOS is one of most researched topic in recent times. Lots of valuable study and thesis is present where different filtering method, several augmentation process, and deep neural network has been utilized. This proposed research focus on the simpler workflow to avoid complexity, because complexity comes with spending more time on deploying models. As this research take place in medical domain, getting ample amount of medical images were very challenging task from the beginning. After gathering data from open source, median filter is implemented to cancel the noise in the images because in medical domain capturing device always produce certain noise in the medical images. This research successfully utilized a median filter where noise was removed from the images but quality of the image drops a bit. Lastly, Transfer learning models (ResNet50, VGG16) are employed along with ImageDataGenrator augmentation process. ResNet50 shows much controlled and balanced prediction compared to VGG16 model. Though accuracy of VGG16 (81%) is much higher that ResNet50 model (52%) but sensitivity drops very low for VGG16 (12%) compared to ResNet50 model (43%), and the same goes for Specificity, that proves ResNet50 model has the better ability to handle both positive and negative instances.

The future scope for this project will be focused on gathering more high quality ultrasound images which will enable the median filter to remove the noise without compromising pixel-to-pixel image quality. As the pre-trained models are utilized here, more automated hyperparameter tuning process such as Grid Search, Random Search, Keras Tuner from TensorFlow can be implemented to get the best combination rather than doing hit and trial method because choosing the best hyperparameters can help to enhance the performance of transfer learning models. In the deployment phase, graphical interface can be made to make it more visually responsive because this research is predicting infected and non infected PCOS after feeding the file name from outside, proper graphical interface can set up for future scope where image upload and prediction can be done within very short amount of time.

References

Abbas, A., Abdelsamea, M. M. and Gaber, M. M. (2020). Detrac: Transfer learning of class decomposed medical images in convolutional neural networks, *IEEE Access* **8**: 74901–74913.

Bharathi, R., Swetha, S., Neerajaa, J., Madhavica, J., Varsha, J., Moorthy, D., Rekha, Sachan, Ramya, S. and Balasundaram, U. (2017). An epidemiological survey: Effect of predisposing factors for PCOS in Indian urban and rural population, *Middle East Fertility Society Journal* **22**(4): 313–316.

Brindha, P. and Rajalaxmi, R. (2023). Comparative study of cnn and transfer learning techniques in the classification of pco ultra sound images, *2023 International Conference on Networking and Communications (ICNWC)* pp. 1–6.

Chang, J., Yu, J., Han, T., jae Chang, H. and Park, E. (2017). A method for classifying medical images using transfer learning: A pilot study on histopathology of breast cancer, *2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom)* pp. 1–4.

Choubey, S. B., Choubey, A., Nandan, D. and Mahajan, A. (2021). Polycystic ovarian syndrome detection by using two-stage image denoising, *Traitement du Signal* **38**(4): 1217–1227.

Do, S., Song, K. D. and Chung, J. W. (2020). Basics of Deep Learning: A Radiologist's Guide to Understanding published radiology articles on deep learning, *Korean Journal of Radiology* **21**(1): 33.

Goel, A., Ahmad, M. O. and Swamy, M. (2019). Design of a 2D Median Filter with a High Throughput FPGA Implementation, *2019 IEEE 62nd International Midwest Symposium on Circuits and Systems (MWSCAS)* .

Hassan, M. M. and Mirza, T. (2020). Comparative analysis of machine learning Algorithms in diagnosis of Polycystic Ovarian Syndrome, *International journal of computer applications* **175**(17): 42–53.

Hijab, A., Rushdi, M. A., Gomaa, M. M. and Eldeib, A. (2019). Breast cancer classification in ultrasound images using transfer learning, *2019 Fifth International Conference on Advances in Biomedical Engineering (ICABME)* pp. 1–4.

Hosain, A. S., Mehedi, M. H. K. and Kabir, I. E. (2022). Pconet: A convolutional neural network architecture to detect polycystic ovary syndrome (pcos) from ovarian ultrasound images, *2022 International Conference on Engineering and Emerging Technologies (ICEET)* .

Jasim, M. K., Najm, R., Kanan, E. H., Alfaar, H. E. and Otair, M. (2019). Image noise removal techniques : A comparative analysis, *International journal of science and applied information technology* **8**(6): 24–29.

Morang, D., Chasta, P. and Chandrul, K. K. (2019). A review on "polycystic overy syndrom pcos", *International Journal of Trend in Scientific Research and Development* **Volume-3**(Issue-4): 60–66.

Motwani, M., Gadiya, M. C., Motwani, R. and Harris, F. (2017). Survey of image denoising techniques, *International journal of advance engineering research development* **4**(7).

Nattadet, C. (2018). Image Filters with Python and OpenCV.
URL: *https://medium.com/nattadet-c/image-filters-41c23f09c600*

Neuzil, A. (2014). What is polycycstic ovary syndrome (pcos)?
URL: *https://www.austinfitmagazine.com/November-2014/what-is-polycycstic-ovarysyndrome-pcos*

Omer, A. A., Hassan, O. I., Ahmed, A. I. and Abdelrahman, A. (2018). Denoising CT Images using Median based Filters: a Review, *2018 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)* .

Vijayan, S. and Venkatachalam, R. (2021). Comparison of filters (dwt, ksvd) and edge detection methods (canny and sobel) on medical images., *Turkish Online Journal of Qualitative Inquiry* **12**(7): 3790–3797.

Wieclawek, W. W. and Pietka, E. E. (2019). Granular filter in medical image noise suppression and edge preservation, *Biocybernetics and Biomedical Engineering* **39**(1): 1–16.

Yang, T., Zhang, C., Ma, Y. and Mao, Y. (2015). Knowledge discovery in databases based on deep neural networks, *2015 IEEE 10th Conference on Industrial Electronics and Applications (ICIEA)* .

Zheng, Y., Yang, C. and , (2018). Breast cancer screening using convolutional neural network and follow-up digital mammography, *Computational Imaging III* .