Enhancing PCOS Detection with SRGAN-Generated Synthetic Images and CNN Models

MSc Research Project
Data Analytics

Ridima Chetan Tambde
Student ID: X22209557

School of Computing
National College of Ireland

Supervisor: Abid Yaqoob

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Ridima Chetan Tambde ……. ……………………………………………………………………………………………
Student ID:	X22209557 ……………………………………………………………………………………………..……
Programme:	MSc in Data Analytics …………………………………………………………… **Year:** 2023 - 2024 …………………………..
Module:	MSc Research Project …………………………………………………………………………………………….………
Supervisor:	Abid Yaqoob …………………………………………………………………………………………….………
Submission Due Date:	12/08/2024 …………………………………………………………………………………………….………
Project Title:	Enhancing PCOS Detection with SRGAN-Generated Synthetic Images and CNN Models …………………………………………………………………………………………….………
Word Count:	7609 ……………………………………… **Page Count** 24 …………………………………………….……..

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: Ridima Chetan Tambde ……………………………………………………………………………………………………

Date: 12th August 2024 ……………………………………………………………………………………………………

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):	

AI Acknowledgement Supplement

Subject Name

Thesis

Your Name/Student Number	Course	Date
X22209557/Ridima Tambde	MSc in Data Analytics	12/08/2024

This section is a supplement to the main assignment, to be used if AI was used in any capacity in the creation of your assignment; if you have queries about how to do this, please contact your lecturer. For an example of how to fill these sections out, please click here.

AI Acknowledgment

This section acknowledges the AI tools that were utilized in the process of completing this assignment.

Tool Name	Brief Description	Link to tool
Chat Gpt	This tool was used for the understanding of SRGAN architectural components.	https://chat.openai.com/

Description of AI Usage

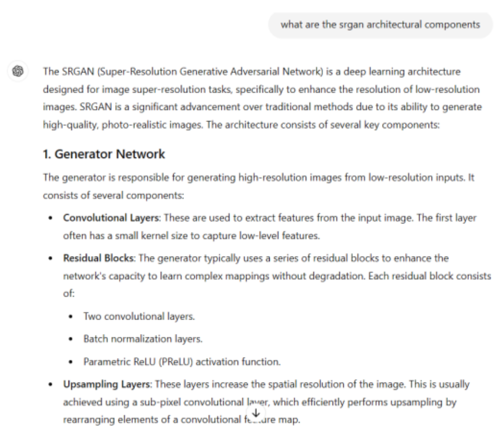
This section provides a more detailed description of how the AI tools were used in the assignment. It includes information about the prompts given to the AI tool, the responses received, and how these responses were utilized or modified in the assignment. **One table should be used for each tool used**.

Chat Gpt	
This tool was used to understand about the architectural layers and components	
Prompt: "what are the srgan architectural components"	The SRGAN (Super-Resolution Generative Adversarial Network) is a deep learning architecture

Evidence of AI Usage

This section includes evidence of significant prompts and responses used or generated through the AI tool. It should provide a clear understanding of the extent to which the AI tool was used in the assignment. Evidence may be attached via screenshots or text.

Additional Evidence



what are the srgan architectural components

The SRGAN (Super-Resolution Generative Adversarial Network) is a deep learning architecture designed for image super-resolution tasks, specifically to enhance the resolution of low-resolution images. SRGAN is a significant advancement over traditional methods due to its ability to generate high-quality, photo-realistic images. The architecture consists of several key components:

1. Generator Network

The generator is responsible for generating high-resolution images from low-resolution inputs. It consists of several components:

- **Convolutional Layers**: These are used to extract features from the input image. The first layer often has a small kernel size to capture low-level features.
- **Residual Blocks**: The generator typically uses a series of residual blocks to enhance the network's capacity to learn complex mappings without degradation. Each residual block consists of:
  - Two convolutional layers.
  - Batch normalization layers.
  - Parametric ReLU (PReLU) activation function.
- **Upsampling Layers**: These layers increase the spatial resolution of the image. This is usually achieved using a sub-pixel convolutional layer, which efficiently performs upsampling by rearranging elements of a convolutional feature map.

Enhancing PCOS Detection with SRGAN-Generated Synthetic Images and CNN Models

Ridima Tambde
X22209557

Abstract

Polycystic Ovary Syndrome (PCOS) is a common hormonal disorder affecting women throughout the world and is detected using ultrasound scans. Due to this accurate diagnosis crucial for effective treatment, is often hindered due to the limited availability of data. For building reliable models, high-quality data is needed and with recent advancements in artificial intelligence, a concept called data generation where synthetic images are generated, has shown promising results. This study attempts the use of Super-Resolution Generative Adversarial Networks (SRGAN) to create synthetic images from existing ones. These synthetic images, along with the original ones, are used to train and test various convolutional neural network (CNN) models, including NasNetMobile, Resnet-152, and Xception. Additionally, hybrid models combining all 3 CNN models with CatBoost are developed and evaluated. The SRGAN architecture is fine-tuned here till good images are obtained and the effectiveness is analyzed to determine their impact on diagnostic performance. Therefore, this research involves a comparison of the classification results from both original images and generated images, thus helping to understand if synthetic data influences the accuracy and reliability of diagnostic models with evaluation metrics like accuracy, precision, recall, F1 score, and the confusion matrix for understanding if any misclassifications. Thus, the study concludes by identifying the most effective model combinations and providing valuable insights for future research in medical imaging.

1 Introduction

1.1 Background

The 21st Century has given rise to advancements in technology and medicine improving the chances of early diagnosis of diseases and the treatment of patients. However, due to these advancements, various ailments persist and impact millions of people from the global population and Polycystic Ovary Syndrome (PCOS) is one among them. PCOD is a very widespread pathology in the female population of the reproductive age, and it appears when ovaries secrete androgens in excessive amounts, which leads to hormonal disruption. This disrupts normal growth and ovulation causing infertility, weight gain, acne and excessive hair growth (Teede, Deeks and Moran, 2010). PCOS affects women's health in the long run as they are susceptible to cardiovascular diseases, type 2 diabetes, and obesity (Kakoly *et al.*, 2019) and low self-esteem among women due to these hormonal changes. Here, early diagnosis takes a central role, although 70% of women with PCOS have not been diagnosed (Wolf *et al.*, 2018). This points to the fact that there is a need to enhance the diagnostic processes. A cross-sectional study conducted with participants at the IEEE conference highlighted the current diagnostic approaches's inadequacies and possibilities for improvement through imaging and AI (Liu *et al.,* 2019).

1.2 Motivation

One significant challenge in medical imaging is the scarcity of high-quality data due to privacy concerns and limited availability, which obstructs the development of AI models. Many studies have used limited datasets such as Suha and Islam (2022) utilized only 594 ultrasound images, and Alamoudi *et al.* (2023) emphasized the need to increase dataset for better model performance and diversity. By generating synthetic images that match with real images, it is possible to increase data and enhance the model performance (Frid-Adar *et al.,* 2018). This research therefore aims to address this gap by employing synthetic advanced data generation technique called Super-Resolution Generative Adversarial Networks (SRGANs) to generate high-resolution synthetic ultrasound images of ovaries. SRGAN are generative models that are capable of enhancing the resolution and details of images, making them reliable for detection (Ledig *et al.,* 2017).

The main purpose of this research is to combine the images generated using SRGAN with the original dataset and evaluate the results using standalone models like NasNetMobile, ResNet152, and Xception. Hybrid models are also employed in this research, where features are extracted from the mentioned CNN models and are further analyzed using machine learning algorithms like CatBoost. By leveraging the strengths of both deep learning and traditional machine learning techniques, this study aims to improve the accuracy of PCOS diagnosis. The SRGAN architecture was fine-tuned involving multiple variations and adjustments to get better details in the image similar to real ones. For this research, the generated data from the last three epochs out of 100 epochs of SRGAN training which produced the best results based on visual inspection is utilized. The original dataset consisted of 1986 images out of which 1588 were training images that were used to train SRGAN and a combined dataset was formed containing original data and the generated data (4764 x 3) resulting in 6352 images in the combined dataset and this combined data was used for classification. The final results are then verified with the results obtained from just the original dataset to understand if the addition of SRGAN made an impact.

The research is based on the research question:

Can the integration of SRGAN-generated synthetic images with real ultrasound images enhance the detection accuracy of CNN-based models, such as NasNetMobile, ResNet-152 and XceptionNet, as well as hybrid models combining CNNs with Catboost for PCOS?

1.3 Research objectives

1. Generating images using SRGAN and assessing the performance in classification.
2. To assess the performance of standalone CNN models (NasNetMobile, XceptionNet, ResNet152) and hybrid models that combine the mentioned CNN models with Catboost on the generated synthetic dataset.
3. To compare the performance of standalone CNN models and hybrid models to know which one is effective.

4. To assess the results of the synthetic data with the original dataset to understand the impact of SRGAN-generated data on the classification.

The report includes Section 1 as Introduction followed by Section 2 as Related Works that reviews existing research, identifies shortcomings, and presents the novel approach in this research. Section 3 is Research Methodology that explains the CRISP-DM framework and the use of SRGAN for data generation and CNN for classification. Section 4 is Design Specification that explains the architecture of SRGAN and CNN models. Section 5 Implementation describes the setup and configuration and Evaluation assesses the model performance. Finally, Section 6 Conclusion and Findings summarizes the project and provides future direction.

2 Related Work

This section reviews recent studies focusing on their methodologies, results, positive impact, and limitations, highlighting the potential challenges faced in leveraging AI for diagnosis.

2.1 PCOS Classification using CNN Models

Hosain *et al.* (2022) introduced PCONet, a custom-designed CNN model for classifying PCOS using two datasets: Dataset A with 1924 training images and 1932 test images, and Dataset B with 339 images. After preprocessing steps like resizing, normalizing, and augmentation, PCONet achieved an accuracy of 98.12%, outperforming fine-tuned InceptionV3 model. However, the train and test sets had identical images, which might contribute to high performance. Similarly, Srivastav, Guleria and Sharma (2024) used a fine-tuned VGG16 model on the same dataset as mentioned in the study by Hosain *et al.* (2022). Their method involved segmentation using techniques like thresholding and edge recognition, achieving an accuracy of 99.4%, but no justification was provided for high results raising concerns for overfitting. A study by Kumar *et al.* (2023) employed MobileNet architecture on a total of 1924 ultrasound images which provided effective outcomes. The author mentions certain limitations like heavy reliance on the quality of ultrasound images, as it can affect the model performance, hence recommends the need to enhance image quality through creating synthetic data. Chitra *et al.* (2023) proposed a hybrid approach using an ensemble of pre-trained models like AlexNet, VGG16, and Inception V3 to classify PCOS. They enhanced the images with segmentation to remove and enhance relevant features, followed by data augmentation. This hybrid model achieved an accuracy of 87%, outperforming the individual models. The study highlights the limitation of decline in performance with increase in data. In the study by Subramani, Rarichan and Chaithra (2023), the authors developed a deep-learning based classification for predicting PCOS on 400 ultrasound images taken from a hospital. The study compared AlexNet, ResNet50, VGG16, and YOLOv5 for feature extraction and classification with YOLOv5 achieving highest accuracy of 99.8%. The positive aspects of this research include the high accuracy, however, the limited dataset can be a problem in terms of diversity. Ravishankar *et al.* (2023) proposed a method combining fuzzy rules with CNN to detect PCOS on 440 ultrasound images obtained from SRM Medical Science Hospital. The methodology involved preprocessing to 64x64 size as fuzzy is computationally intensive, followed by feature extraction by CNN, and then fuzzy rules were integrated. This model achieved an accuracy of 98.37% with metrics such as precision, recall and F1 indicating high performance. The limitations highlighted were integrating fuzzy logic with CNN can be computationally intensive and resizing to 64x64 can cause feature loss.

2.2 PCOS Classification using Machine learning models

Sumathi *et al.* (2023) proposed a methodology for detecting PCOS using a dataset of 1918 augmented images from 25 hospital images. Their process included image acquisition, enhancement with Histogram Equalisation, segmentation, feature extraction using Gray Level Co-occurrence Matrix (GLCM) and classification. Darknet-19, AlexNet, SqueezeNet, and SVM were used, with SVM achieving the lowest accuracy of 84% and Darknet with 99%. The research stresses the need for more data. In contrast, Purnama *et al.* (2015) used two datasets of Dataset A with 275 ovary images based on mean texture and Dataset B with 399 images based on other statistical measures. Features were extracted using Gabor wavelet method, and classification was performed using Neural Network-Learning Vector Quantization (LVQ), K-Nearest Neighbor (KNN), and Support Vector Machine (SVM). SVM achieved the highest accuracy of 82.55% on Dataset A, and KNN with 78.81% on Dataset B. They highlighted the need for fine-tuning and refining the models for better results. In contrast, Suha and Islam (2022) utilized 594 ultrasound images of ovaries and used a hybrid model with features extracted by VGG16 and combined with a stacking ensemble of 5 traditional classifiers (Logistic Regression, SVM, Decision Tree, KNN, Naïve Bayes) and XGBoost as the meta-learner. This approach achieved the highest accuracy of 99.89% on limited dataset. While all studies used image data, the study by Deshpande and Wakankar (2014) involved using a combination of image data and clinical data, containing total of 20 ultrasound images, and dataset based on clinical parameters was used. Feature extraction was performed using a multiscale morphological approach, followed by segmentation and Canny edge detection and classification was done using SVM with a 95% accuracy. The author highlighted expanding the dataset in the future.

2.3 Generative Adversarial Networks (GANs)

Frid-Adar et al. (2018) applied both traditional data augmentation and synthetic data generation using Generative Adversarial Networks to classify liver lesions from CT images. The study used a small dataset of 182 images and applied traditional augmentation which resulted in 30,000 images which were further classified by CNN. The results were compared with classification results obtained by using synthetic data generated through DCGAN (Deep Convolutional GAN), a GAN architecture. The accuracy improved from 78.6% with traditional augmentation to 85.7% with the inclusion of synthetic data. On the contrary, Ahmad *et al.* (2022) presented SRGAN for improving image quality on four datasets including skin, retinal, cancer and cardiac ultrasound images forming a total dataset of 1400 images. The architecture used initial layers for feature extraction with ResNet34, final layers with progressive upscaling with residual connections and CNN. SRGAN significantly improved image quality, achieving a PSNR of 38.83 dB and SSIM of 0.95. The limitation of this study was the use of Loss L1 function which can reduce errors but miss important features. The study by Liang *et al.* (2022) proposes a methodology using PGAN (Progressive-GAN) based framework to synthesize ultrasound images. The study uses 3 large datasets: lung, hip joint and ovary and the approach employs auxiliary sketch guidance, progressive training strategies, and feature loss to enhance image quality. This approach showed promising results with FID ranging from 0.50 to 0.60 for the datasets, however, a limitation of the study is that the generated high-resolution images can be blurry and contain salt-and-pepper noise. Shashank, Acharya and Sivaraman (2023) used SRGAN with a VGG-19 based loss function to enhance feature details and quality, showing significant improvements but faced challenges with overfitting needing a deeper network. Nandhini, Srinath and Veeramanikandan (2022) combined SRGAN and CNN to detect glaucoma from 680 retinal fundus images that significantly improved accuracy from 75.20% to 96.11% and reduced loss metrics from 32.07% to 13.36%. For CNN, the limitations included reliance on limited and high-quality data as it increased the performance.

The above studies show significant advancement in PCOS detection, but the gap remains. Studies including Kumar et al. (2023) and Suha and Islam (2022) use limited datasets, which can hinder performance and diversity. For PCOS classification, CNN models like NasNetMobile, ResNet152 and Xception, as well as their combination with the machine learning model CatBoost, have not been previously applied to PCOS classification. The use of GANs, particularly SRGAN is also unexplored in this domain. As discussed in Section 2.3, SRGAN can generate high-resolution synthetic images. This research proposes a novel approach by generating images with SRGAN and combining them with original data to enhance diagnostic accuracy using standalone CNN and hybrid models

3 Research Methodology

The methodology of this research is structured according to the CRISP-DM framework, which is considered a best practice for data-driven projects due to its flexible process model that ensures a systematic approach to problem-solving. Figure 1 illustrates the detailed process flow for this research.

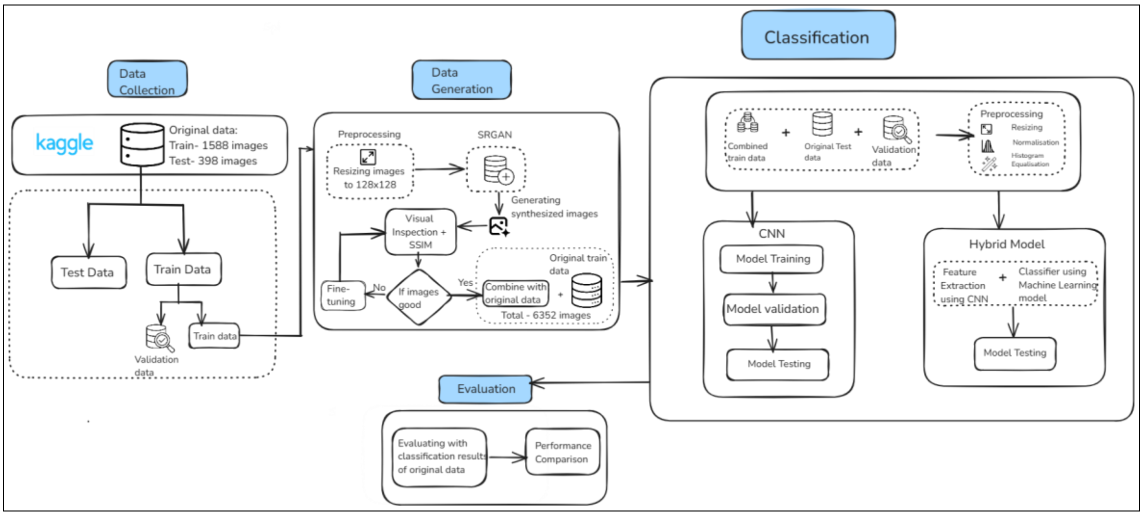


Figure 1: Methodology Diagram.

3.1 Business Understanding

The overall aim of this study is to make a diagnosis of PCOS more accurate by using advanced AI techniques. Diagnosis of PCOS can be improved by generating synthetic data which will lead to better results for patients, and it will bring down the cost by avoiding relying on large volumes of data, lowering the cost, and increasing the speed of the development process.

3.2 Data Understanding

The dataset[1] utilized in this study is taken from Kaggle consisting of 1588 training images and 398 test images of ultrasound scans of ovaries, labeled as PCOS and Normal.

3.3 Data Preparation

Initially, each image's mode was checked to determine if it was stored in RGB or Grayscale format using the PIL library of Python. The images appeared grayscale after this check.

SRGAN Pre-processing: The images were loaded and resized to 128x128 pixels to standardize the input size for the SRGAN model and then the images were normalized to the range [-1, 1] in order to transform the images into numpy arrays for further processing. The preprocessing steps for SRGAN were also highlighted in the study by Takano and Alaghband (2019).

Classification Pre-processing: This involved preprocessing the images to make sure they are suitable for CNN and hybrid models. The preprocessing was divided into two parts.

1. Pre-processing for classification using combined data: The original data, and SRGAN-generated data involved the same pre-processing steps. Both the datasets were merged to form a combined dataset and all the preprocessing was applied to the combined dataset. The images were resized to 224x224 pixels for models NasNetMobile and ResNet-152 and 299x299 for Xception. The normalization range [0,1] was the same for all three models. Histogram equalization was also applied to both images to enhance the features as this technique is effective for ultrasound images by revealing more details which are often not prominent enough due to its low contrast and noise.

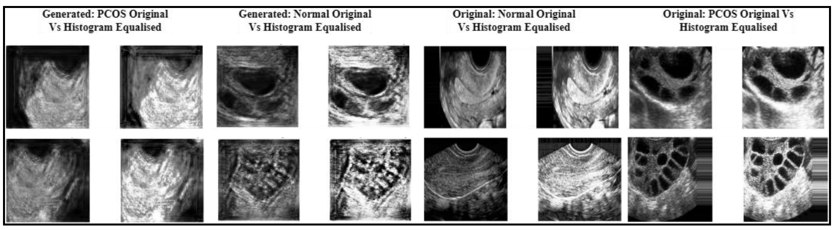


Figure 2: Original Training and Generated Data - Histogram Equalization.

Figure 2 illustrates images before and after histogram equalization on the combined dataset to demonstrate the enhancement of features.

2. Training and Testing: After data preparation, the dataset of 1588 images was split into 80% for training and 20% for validation, and the test set of 398 images remain unchanged for

[1] https://www.kaggle.com/datasets/bharatsh001/pcos-cleaned-and-splitted

final evaluation. The CNN models (NasNetMobile, ResNet152, Xception) and their combination with Machine learning classifier (CatBoost) were trained for 100 epochs using the Adam Optimizer and binary cross-entropy loss functions. Validation data was used during training to fine-tune the models and prevent overfitting and the test set was used to evaluate the model's performance using metrics such as accuracy, precision, recall, F1-score and confusion matrix.

3. Data Generation using SRGAN: In this phase, SRGAN creates high-resolution synthetic images from the original training dataset. This method is preferred because GANs generate more realistic and diverse data. The SRGAN architecture includes a generator and discriminator, which are both convolutional networks to help the model effectively capture complex features (Ledig *et al.,* 2017). With further fine-tuning, the generator employed residual blocks to improve the image quality making sure the images remain detailed (Lim *et al.,* 2017). Adam optimizer with different learning rates was used to help the model learn faster and handle complexities better. This was adjusted for every fine-tuning as the model was built depending on the results. There were two losses used where adversarial loss using binary-cross entropy (BCE) helps the discriminator distinguish between real and fake images and the content loss using Mean Squared Error (MSE) is used to ensure the generated images are structurally similar to the original ones. Additionally, perpetual loss is calculated using a pre-trained VGG19 network to focus on high-level features as this approach is effective for grayscale images by enhancing structure details and textures (Moran *et al.,* 2021). The discriminator is composed of convolutional layers with spectral normalization and Leaky ReLU to help stabilize training and improve performance. The SRGAN model was trained for 100 epochs with a batch size of 32. Images from the last 10 epochs were saved for evaluation, representing the highest quality outputs after extensive training.

3.4 Data Modelling

Data Splitting: The original data was already split into train and test and the train data was used to train SRGAN, which resulted in 4764 images as images generated for the last 3 epochs (97,98,99) from 100 epochs were considered. The original train data was split into validation set with a ratio of 80:20, and the remaining train images were combined with the SRGAN-generated images resulting in 6034 images for classification (1270 remaining train images + 4764 SRGAN-generated images).

CNN Models: Three CNN models (NasNetMobile, ResNet152, XceptionNet) were trained on the combined dataset (original + generated) and validation. The models were initialized with pre-trained weights from ImageNet, and the top layers were replaced with custom layers suitable for PCOS classification. The common architecture for the CNN models included a base model followed by global average pooling, dense layers with ReLU activation, dropout for regularization, and a final dense layer with softmax activation for classification. Finally, the model was compiled by Adam Optimizer.

Hybrid Models: This was constructed by extracting features using CNN and classifying using a traditional machine model, in this case, Catboost.

3.5 Evaluation

The performance of the models trained on the original dataset was compared with training on the combined dataset to evaluate the effectiveness of the SRGAN-generated synthetic images. The evaluation metrics included accuracy, precision, recall, F1 score, confusion matrix, and Area under curve.

3.6 Deployment

In this study, except for the deployment phase, all other phases are executed.

4 Design Specification

4.1 Super-Resolution Generative Adversarial Networks (SRGANs)

This section describes the use of SRGAN to generate additional images for more accurate PCOS detection. The SRGAN consists of 2 main components: the generator and the discriminator. The generator generates detailed images from low-quality input, and the discriminator on the other hand compares the generated images to the real ones. During development of architecture, both components were modified and fine-tuned across several variations to enhance image quality. Pre-processing for all variations was done by resizing them to 128x128 pixels and normalizing them within [-1, 1].

Variation 1: This variation shows a basic framework of a generative adversarial network. Resized images of 128x128 pixels are down-sampled to 64x64 pixels before giving it as input into the generator. Even though the images get downsampled, the pre-processing to resize them to 64x64 is because if the images were initially resized to 128x128 instead of directly resizing them to 64x64, essential details or features can be lost. The downsampling of the image before feeding it into the generator is standard process for SRGAN as it makes the neural network concentrate on extracting and learning critical features from a compressed representation of the image, in a way teaching the network how to reconstruct finer details from less data[2].

Generator - The generator starts by using a 9x9 convolutional layer with 64 filters to extract essential features from the downsampled images and after feature extraction the image is upsampled back to 128x128, which increases the size and adds more detail (Maqsood *et al.*, 2021). After upsampling, an additional 3x3 convolutional layers further refine the texture and structure of images. Finally, the output layer uses 9x9 convolutional filter with tanh activation to give output of the final high-resolution images.

Discriminator - A 128x128 image generated as output from the generator is used as input here and this input will contain both generated and original images (128x128). The initial layers use 64 filters of 3x3 with a LeakyReLU activation function to introduce non-linearity to help capture complex patterns (Jia *et al.*, 2019). There is an addition of more layers with more filters to refine the feature extraction process as the architecture progresses. Some layers are configured with stride=2 to reduce spatial dimensions of feature maps obtained

[2] https://medium.com/analytics-vidhya/super-resolution-gan-srgan-5e10438aec0c

through other convolutional layers, in order to focus on larger features like shapes and textures instead of pixel-level details. Batch normalization is applied after the convolutional layers have processed the image to stabilize the training and help the discriminator learn more efficiently. After the convolutional layers have processed the image, the image data is flattened into a one-dimensional vector for Dense layer, which combines the extracted features and forms an understanding of the image content. The final layer is an output layer with sigmoid activation suitable for binary classification, which enables the discriminator to determine whether an image is real or generated. Figure 3 illustrates the architecture diagram for Variation 1.

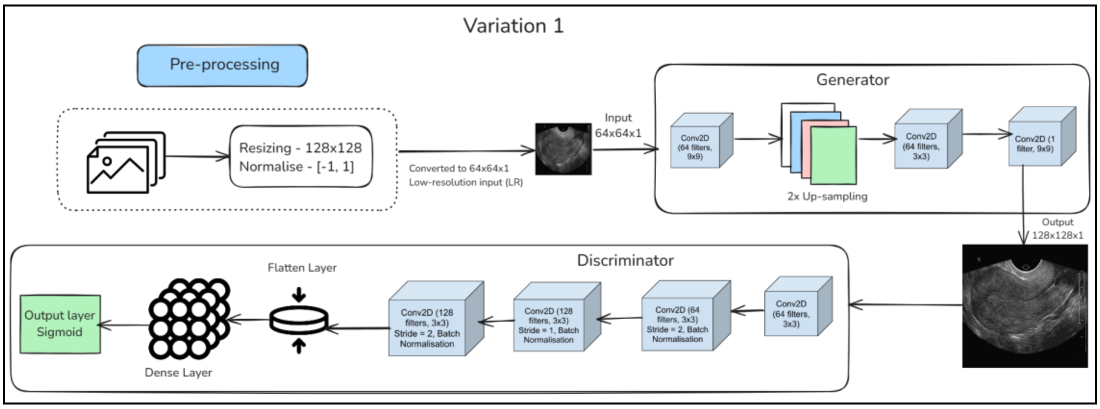


Figure 3: SRGAN architecture diagram Variation 1.

Training Loop - A batch of original and generated images are randomly selected and passed to the discriminator, which evaluates them as real and fake using the architecture with the layers mentioned in Figure 3. During this process, the generator receives feedback in the form of loss (d_loss for the discriminator and g_loss for the generator), which is calculated using Binary Cross-Entropy (BCE), a loss function that measures the difference between actual and predicted values. This feedback helps the generator improve its ability to create realistic images. Through backpropagation, the generator fine-tunes its parameters to minimize g_loss and fool the discriminator. The training progresses with each batch at 0.0001 learning rate for both generator and discriminator, controlling the rate at which the model learns. Figure 4 demonstrates this training loop for Variation 1.

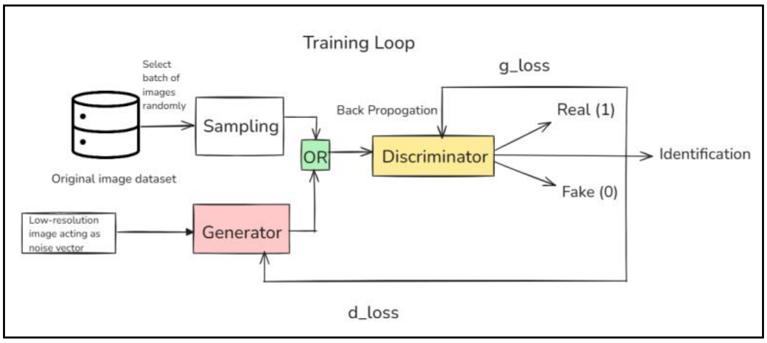


Figure 4: Training loop.

<u>**Variation 2:**</u> The architectural setup is majorly similar to Variation 1 including the preprocessing, but there are several enhancements. Variation 2 introduces residual blocks, which are crucial to help maintain the image quality and details as they pass through multiple layers of the network.

Generator **-** Residual blocks allow the network to learn and apply small changes on top of the original image data rather than having to recreate the image details from scratch at each layer, thus preventing loss of important details which can happen if crucial details are lost as it passes through many layers. This addition is important in architectures like SRGAN as it resolves the issue of loss of information. The residual block in this variation consists of two convolutional layers, followed by batch normalization to enhance the model's ability to capture and maintain essential details and PreLU activation to allow the network to continue learning from images. Figure 5 illustrates the residual block architecture for this variation.

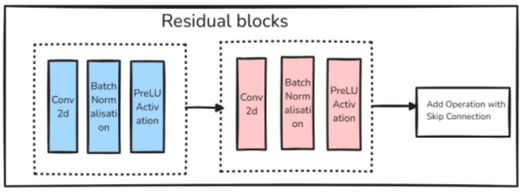


Figure 5: Residual blocks layers – Variation 2.

Discriminator **–** In variation 2, the discriminator is improved by adding Spectral Normalisation to all convolutional layers, because in the context of SRGAN, without this the discriminator might become too powerful, making it hard for the generator to improve as the generator would easily identify generated images as fake and this could stop the feedback to the generator and stopping the progress, thus spectral normalization will help to bring that balanced learning (Miyato *et al.*, 2018). In this variation, 2 more convolutional layers are also added with filters, increasing the discriminator's ability.

Training loop – The training loop in Variation 2 follows the same process as the training loop in Variation 1 except for the learning rate which is increased to 0.0002 allowing the model to learn more quickly.

Variation 3: While the preprocessing remains the same across all the variations, there are certain adjustments made in the generator and discriminator for this case.

Generator – Similar to variation 2, the generator here also contains residual blocks, but the number of residual blocks is increased from 5 to 15 allowing the network to process and enhance the details through deeper layers, in a way making sure the texture of the image is also improved.

Discriminator – The architecture is the same as the one for variation 2 and similarly, spectral normalization is maintained across all the layers. Figure 6 illustrates the architecture diagram for Variation 2, Variation 3 and Variation 4.

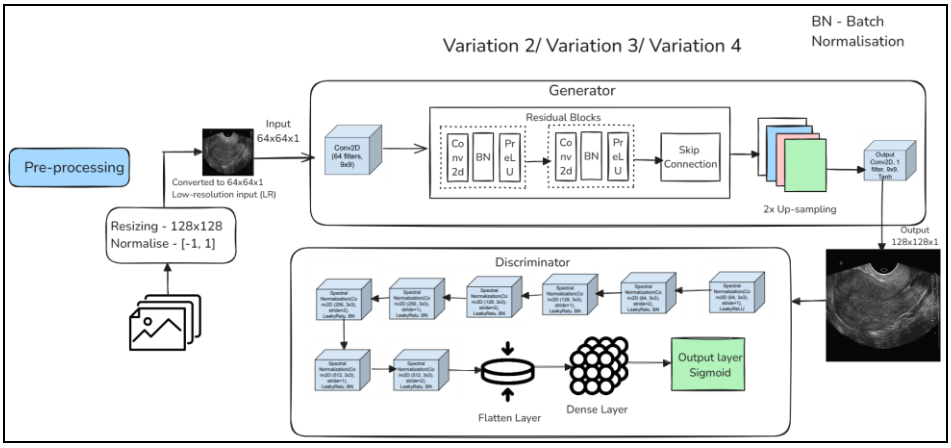


Figure 6: Architecture Diagram for Variation 2/Variation 3/Variation 4.

Training Loop – The training loop process here is similar to the training loop for variations 1 and 2, however VGG-19 for feature extraction is introduced. Both generated and synthetic images are passed to VGG-19 which extracts high-level features from the images which is then used to compute the perpetual loss. This perpetual loss evaluates how similar both (synthetic + original) images look beyond just pixel-to-pixel comparison. For normal images it will detect smooth textures without cysts, and for PCOS it will learn to detect complex textures with multiple cysts. This perpetual loss is then combined with the discriminator loss to create a total loss and provide feedback to generator. Figure 7 illustrates the training loop for Variations 3 and 4.

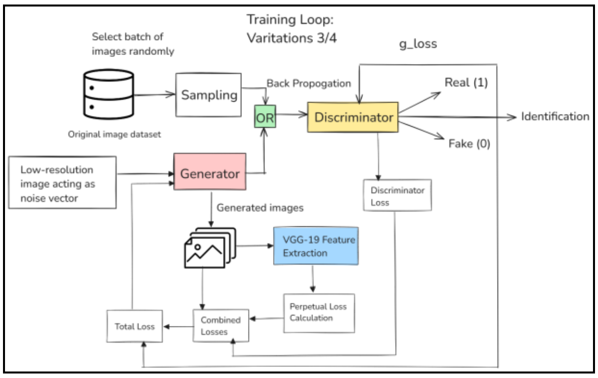


Figure 7: Training loop for Variations 3 & 4.

Variation 4: The architecture for variation 4 remains the same for the generator and discriminator with residual blocks increased from 15 to 16 to capture more depth in features, but there are some enhancements made in the training loop.

Training Loop - In variation 4, the training loop incorporates sophisticated mechanisms to optimize the learning ability. The learning rate scheduler with exponential decay is employed which methodically lowers the learning rate as training progresses as it allows the neural network to make finer adjustments useful in the later stages of training to prevent exceeding the minimal loss values. Additionally, gradient clipping is used to keep the changes made during learning from becoming too large which causes instability (Gulrajani *et al.,* 2017). Label smoothing of 0.9 is also applied to make sure the discriminator does not get too confident with its predictions, promoting a balanced training environment.

4.2 Convolutional Neural Network (CNN)

In this project, 3 standalone CNN pre-trained models NasNetMobile, ResNet152, and Xception are utilized. These models help classify the images by processing them through several layers with ResNet152 offering deep residual networks helping capture intricate details and preventing loss of information and Xception with its depthwise separable convolutions capturing local features more distinctly and NasNetMobile for its lightweight and efficient architecture. Each model starts with a pre-trained based on ImageNet to quickly recognize general image features. Images are resized to fit model specifications: 224x224 for NasNetMobile and ResNet152, and 299x299 for Xception. These models come with several built-in layers such as Batch Normalization which ensures that the input data to each layer is standardized and ReLU activation which introduces non-linearity, helping the model learn complex patterns. Then, the images are sent to Global Average Pooling (GAP) layer which reduces the number of parameters by focusing on the important features and a dropout layer is then applied to avoid overfitting. Finally, a dense layer with Softmax activation is used to make final classification based on the features extracted. Figure 8 illustrates the architecture diagram of CNN standalone model describing image classification through different layers.

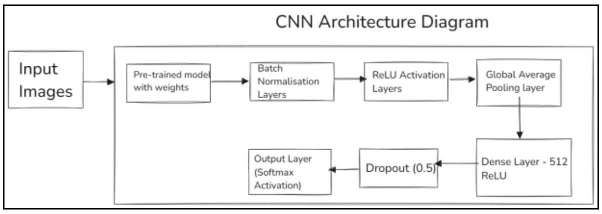


Figure 8: CNN pre-trained models Architecture diagram.

4.3 CNN and Machine Learning Classifier

In this setup, images first go through a pre-trained model. Similar to CNN architecture in Figure 8, with pre-trained networks, batch normalization layer, ReLU activation layer remain consistent along with GAP which is embedded with these pre-trained models. These layers of CNN then process the features at every step and then these processed features are used by the Catboost classifier, a type of machine-learning model that excels in recognizing complex patterns within the image data, making it suitable for classifying. This classifier makes the final decisions, labeling the images based on the features it receives from initial image processing. Figure 9 displays the architecture for CNN + CatBoost classifier architecture.

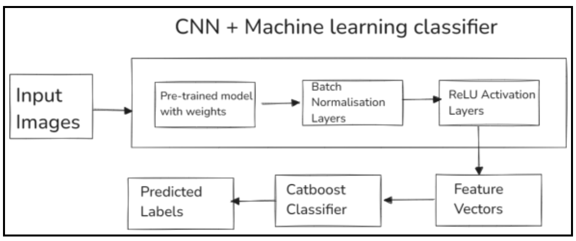


Figure 9: CNN + Catboost classifier Architecture.

5 Implementation

5.1 Setup and Data Configuration

The original data used in this project was already split into train and test. Training consisted of 913 images of normal class and 675 images of PCOS class, test set included 229 normal and 169 PCOS class images. To create more data for enhancing model training and diversification, SRGAN was employed and the generator was trained on the original training dataset on 100 epochs where each epoch produced 1588 images (913 + 675). During the training, images were processed with a batch size of 32 to manage computational load effectively. From the numerous variations and adjustments made to the SRGAN model, the fourth variation generated good images from epoch 80 to epoch 100 but only images from epochs 97, 98, 99 were considered due to computational requirements and based on clarity and visual inspection. This resulted in 4764 SRGAN-generated images. The generated data was then combined with the original training set and the original test set remained unchanged to ensure the model effectively handles unseen data. To balance the large training set and mitigate the risk of overfitting, the original training data was split into an 80:20 ratio to provide a reliable means of fine-tuning the model's parameter through the validation feedback.

This was executed on the Google Colab platform utilizing a TPU v2 with 300GB due to the intensive computational abilities of SRGAN and further classification with 6352 images.

5.2 Pre-processing and Data Transformation

The CNN models in this research like VGG19 for SRGAN, NasNetMobile, ResNet152, Xception for CNN require 3-channel RGB input. To address this, the grayscale images were converted to RGB by duplicating the grayscale channel across all 3 RGB channels using the PIL library to be suitable for CNN.

SRGAN Preprocessing – For SRGAN, the original training dataset was used and the images were resized to 128x128 pixels to maintain uniformity. Then, the images were normalized to [-1, 1] to help the network learn better. Additionally, this normalization was done since 'tanh' function used in the output layer supports [-1,1] normalization.

CNN Preprocessing – Preprocessing was applied to a combined dataset of 6352 images, consisting of 1588 original training images and 4764 generated images from the last three epochs (97, 98, 99). The validation (split from training) and test set were also preprocessed. Images were resized as per the model's input requirements: 224x224 for NasNetMobile, ResNet152 and 299x299 for Xception. Standard normalization of [0,1] range was applied, followed by histogram equalization to enhance contrast and reveal more features. Then, label encoding was used to convert categorical labels into numeric form.

5.3 Implementation: SRGAN Variations

In the process of generating synthetic data, various configurations of the SRGAN model were tested to identify the best-performing architecture. This involved modifying key parameters and components in the generator and discriminator network and Table 1 summarises the key parameters and differences.

Table 1: SRGAN Model Variations and Parameters

Variation	Generator	Discriminator	Learning Rate	Activation Function	Optimizer	Loss Function
1	Basic CNN	Standard layers CNN	Gen: 0.0001, Disc: 0.0001	LeakyReLU	Adam (β1=0.5)	Binary Cross Entropy (BCE)
2	CNN with 5 residual blocks	Spectral Normalisation (SN) + Batch Normalisation (BN)	Gen: 0.0002, Disc: 0.0002	PreLU	Adam (β1=0.5)	Binary Cross Entropy (BCE)
3	CNN with 15 residual blocks	SN + BN	Gen: 0.0002, Disc: 0.0002	PreLU	Adam (β1=0.5)	VGG-19 based perpetual loss
4	CNN with 16 residual blocks	SN + BN	Starts at 0.0001 Decay	PreLU	Adam with gradient clipping (β1=0.5, β2=0.999)	VGG-19 based perpetual loss

5.4 CNN and Hybrid Models

The combined dataset and the validation dataset were used during training and the results were evaluated against the test data. The classification was done with the help of standalone CNN models (NasNetMobile, ResNet512, Xception) and hybrid models (CNN model for feature extraction, and the extracted features were classified using the CatBoost classifier). The models were compiled using Adam Optimizer, known for its ability to handle complex data and adjusting the learning rate automatically. Sparse categorical loss was also used to handle the classification of labeled data. Each model was trained on 10 epochs and the batch size was 32 to optimize memory usage and ensure data samples accurately reflect the overall dataset. Validation was conducted using the 20 percent data split from train data to evaluate model performance and avoid the risk of overfitting.

For Hybrid models, the features were extracted using the same CNN models (NasNetMobile, ResNet512, Xception) and were classified using Catboost classifier, known for its performance with categorical data with training parameters such as learning rate of 0.1, tree depth of 4 and 100 iterations. The 'verbose' parameter was set to 10, providing feedback after every 10th iteration.

6 Evaluation

The effectiveness of image generation using SRGAN and classification using CNN and hybrid models is evaluated to understand the impact made by SRGAN-generated data on limited data. For SRGAN, four variations are analyzed to show improvements in image features using visual inspection and SSIM score, and discriminator vs generator loss graph. For classification performance is evaluated using classification reports, confusion matrix, accuracy, ROC curves, and loss graphs. SSIM refers to the similarity between real and generated images in terms of pixel and structure. Classification metrics include precision (correct predicted values), recall (correct actual values), specificity (correct actual negatives), F1 score (combines precision and recall), and ROC curves (balance between true positives and false negatives).

6.1 SRGAN Variation 1 / Variation 2

Visual Inspection: The generated images are assessed over 100 epochs to observe the progression in image quality. Images displayed are saved in a grid of 25 samples for both classes at every 10th epoch to monitor how the image generation progresses and how much the images can adapt the features of the original dataset. Figure 10 illustrates the display of generated images from SRGAN Variation 1 and Variation 2, showing the evolution of image quality over 100 epochs. These images show significant blurriness and lack of fine detail. For Variation 1, if PCOS class is observed, there is little indications of cystic structures but normal class shows significant blurriness and hence images from these variations are not considered. For Variation 2, images are getting saved every 10th epoch similar to Variation 1. But overall epochs also demonstrate a similar display of images, hence for evaluation images of the last epoch are shown. Both PCOS and normal classes show overwhelming darkness and lack visible detail, making it difficult to distinguish between the two classes. Due to these shortcomings, further metrics like SSIM and generator-discriminator loss graph will not be shown.

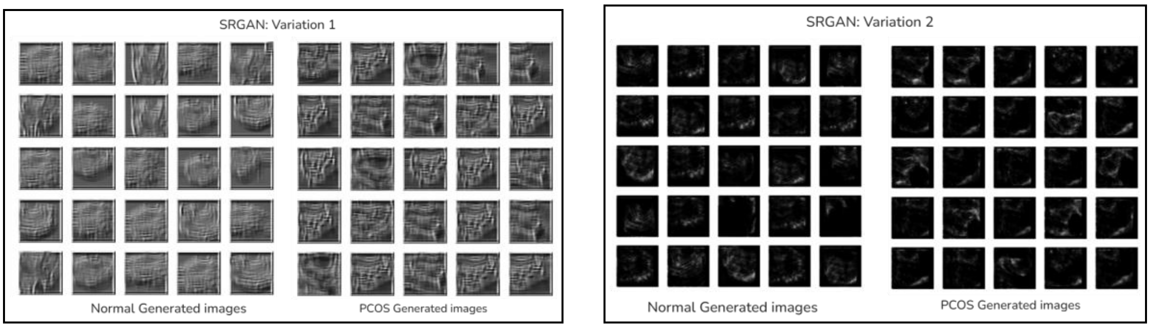


Figure 10: Generated images – Variation 1/ Variation 2.

6.2 Variation 3

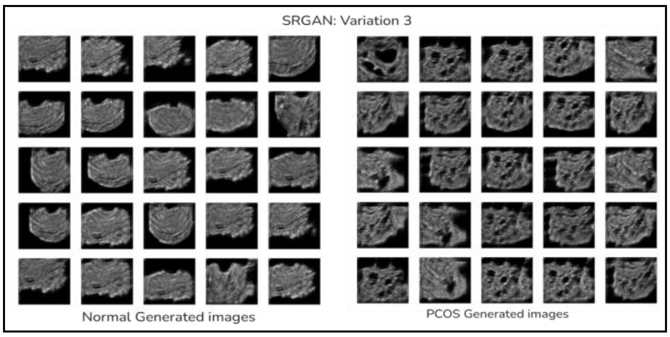


Figure 11: Generated images: Variation 3.

The images for variation 3 as shown in Figure 11 show some differences from the other 2 variations. The normal class shows improved texture and more visible details than variation1, but there is still significant blurriness. For the PCOS class, there's a noticeable enhancement in the visibility of cysts which makes it easily understandable but it still can't be considered to be taken as for both classes, each image does not fully occupy the given space leaving noticeable black space and incomplete usage of the frame, thus further adjustments need to be made and these images cannot be considered for classification.

6.3 SRGAN Variation 4

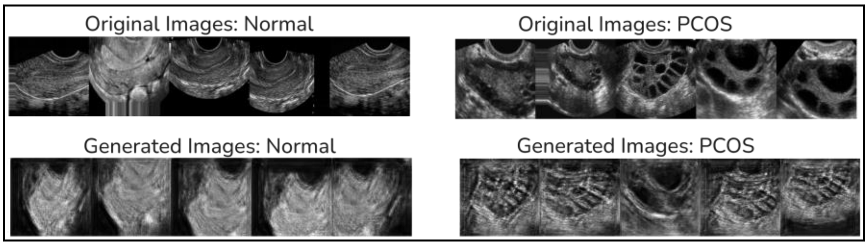


Figure 12: Generated vs Original images – SRGAN Variation 4.

Figure 12 illustrates the comparison of original and generated images from Variation 4 over 80-100 epochs. In SRGAN Variation 4, improvements in image generation were observed between epochs 80 to 90 and as the epochs progresses, it gets better, especially for epochs 97,98,99. The normal generated images are improved in capturing structural elements compared to previous ones, but there is still little blurriness. But, there is a visible effort by the model to replicate the general textures and shapes found in the original image. The PCOS-generated image shows better cystic structures suggesting that the model is learning to distinguish and replicate important features. It is observed that the blurriness is constant but it also somewhat occurs from the original image as they also have a little blurriness factor. A conclusion can be made that after increasing residual blocks from 15 to 16 and introducing a finer controlled exponential decay for the learning rates to decrease slowly, the model's ability to generate stable images improved.

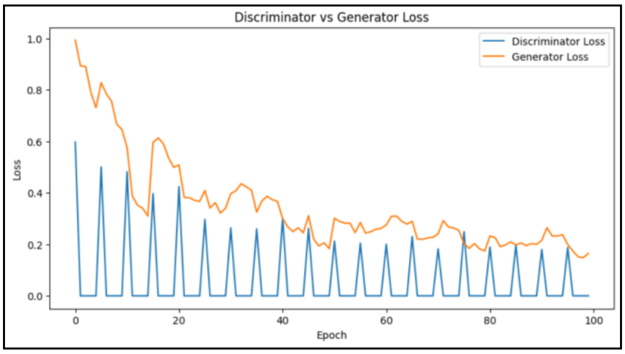


Figure 13: Discriminator vs Generator Loss Graph – SRGAN Variation 4.

In Figure 13 the graph shows discriminator and generator losses across 100 epochs. Initially, both losses experience high fluctuations at the early stages of the model where the model is learning to refine its image generation. As training progresses, the losses decrease and start to stabilize, indicating the generator is producing more accurate images and the discriminator is learning effectively. This decrease in loss value suggests that the model has reached a consistent level of performance, capturing and replicating ultrasound images.

Table 2: SSIM metrics for SRGAN Variation 4

Epoch	SSIM for Normal Class	SSIM for PCOS class
97	0.16	0.16
98	0.11	0.18
99	0.1	0.13

Table 2 shows SSIM metrics for SRGAN Variation 4 across last three epochs with visually good images. While the SSIM values are not very high, indicating structural differences between original and generated images, it is important to note that SSIM evaluates images on a pixel-by-pixel basis, focusing on specific aspects like brightness and structure. Visually, these images from the last 3 epochs show better improvements and capture features, thus considering these images can be useful in analysis, resulting in more diverse data. For classification, only the images from epochs 97, 98, 99 are taken that corresponds to 15% of the total generated data from epochs 80 to 100. This is because even though all the images generated in these epochs are good, taking all of them is not possible due to computational intensity and per epoch 1588 images are generated so in case of 80 to 100 epochs, there will be a total of 31,760 images which will be very high computation and hence only 4764 generated images will be considered.

6.4 CNN Standalone models

This section aims to assess whether synthetic images generated using SRGAN can improve the diagnostic accuracy of CNN in detecting PCOS.

NasNetMobile: This model shows an accuracy of 97%. It identifies all normal cases correctly causing a specificity of 100% and PCOS which is the positive class, has 13 misclassifications with a recall of 92%. The F1 score combining precision and recall is strong for PCOS (0.96) and Normal (0.97) indicating balanced performance.

Table 3: Evaluation Metrics for NasNetMobile

Metric	Normal	PCOS	Overall Metrics
Precision	0.95	1.00	
Recall (Sensitivity)	1.00	0.92	Sensitivity: 0.92
F1-score	0.97	0.96	
Test Loss			0.093
Accuracy			97%
Specificity			1.00
Confusion Matrix	TN: 229, FP: 0	TP: 142, FN: 13	

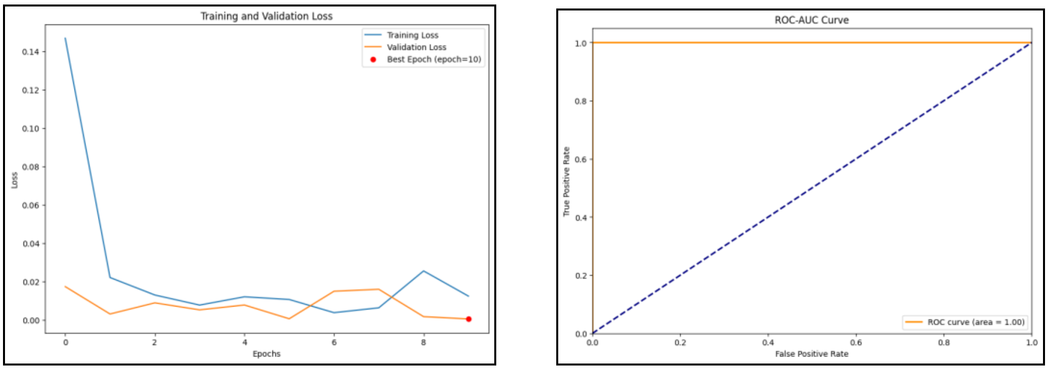


Figure 14: NasNetMobile Training Progress and ROC Analysis.

The training and validation loss graph shown in Figure 14 explains the model stabilized after a few epochs indicating good learning with epoch 10th as the best epoch and ROC-curve of 1.00 confirms the model's ability to distinguish both classes. The model also demonstrates low test loss of 0.093 suggesting precise predictions. Table 3 presents evaluation metrics for NasNetMobile's performance.

ResNet152: This model demonstrates 96% accuracy and a test loss of 0.153. It excels in identifying Normal cases with precision of 93% as 16 PCOS cases were misclassified as Normal but achieves a perfect recall of 100%. For PCOS, precision is 100% because no Normal cases were incorrectly classified as PCOS with a recall of 90%. The F1 scores for PCOS and Normal are 95% and 97% respectively. The training and validation graph in Figure 15 shows the 10th epoch as the best epoch and an ROC-AUC score of 0.99 indicates proper classification. Table 4 presents the evaluation metrics for ResNet152 for performance.

Table 4: Evaluation Metrics - Resnet152

Metric	Normal	PCOS	Overall Metrics
Precision	0.93	1.00	
Recall (Sensitivity)	1.00	0.90	Sensitivity: 0.90
F1-score	0.97	0.95	
Test Loss			0. 153
Accuracy			96%
Specificity			1.00
Confusion Matrix	TN: 229, FP: 0	TP: 139, FN: 16	

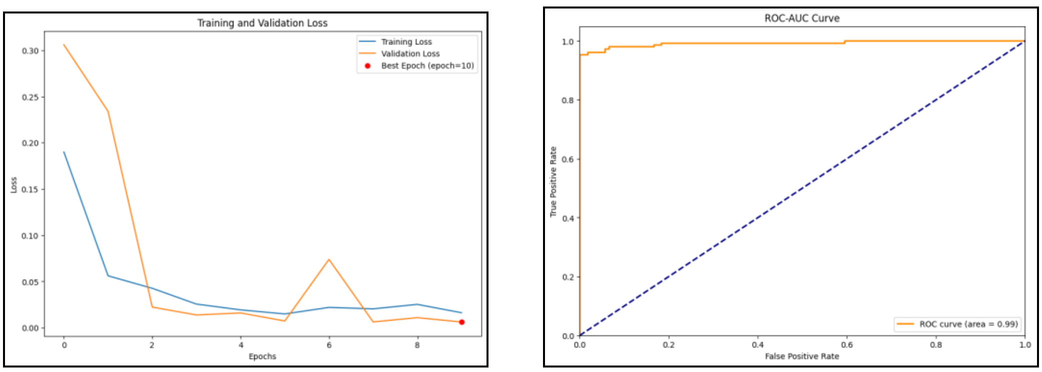


Figure 15: ResNet152 Training Progress and ROC Analysis.

<u>Xception:</u> This model gives an accuracy of 98% and correctly identifies all cases of Normal class but the precision is 97% as 8 PCOS cases were incorrectly classified as Normal, but the recall for Normal is 100%. For PCOS, the precision is 100% as there were no cases where Normal was misclassified as PCOS and the recall was 95% due to 8 missed cases. Table 5 shows the evaluation metrics for Xception indicating the performance.

Table 5: Evaluation Metrics: Xception

Metric	Normal	PCOS	Overall Metrics
Precision	0.97	1.00	
Recall (Sensitivity)	1.00	0.95	Sensitivity: 0. 92
F1-score	0.98	0.97	
Test Loss			0. 1026
Accuracy			98%
Specificity			1.00
Confusion Matrix	TN: 229, FP: 0	TP: 147, FN: 8	

With the high F1 scores of 97% and 98% and the rapid decrease in loss as shown in Figure 16 and the model's nearly perfect ROC-AUC curve, indicated the model's predictions are reliable

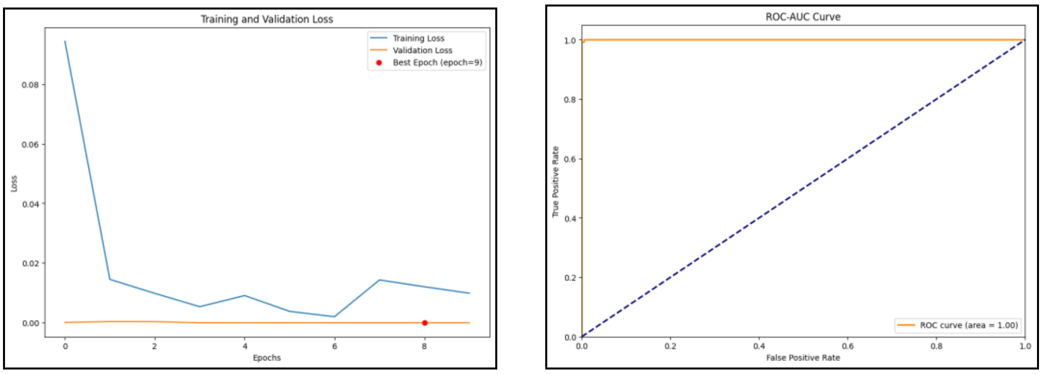


Figure 16: Xception Training Progress and ROC Analysis.

6.5 CNN Hybrid Models

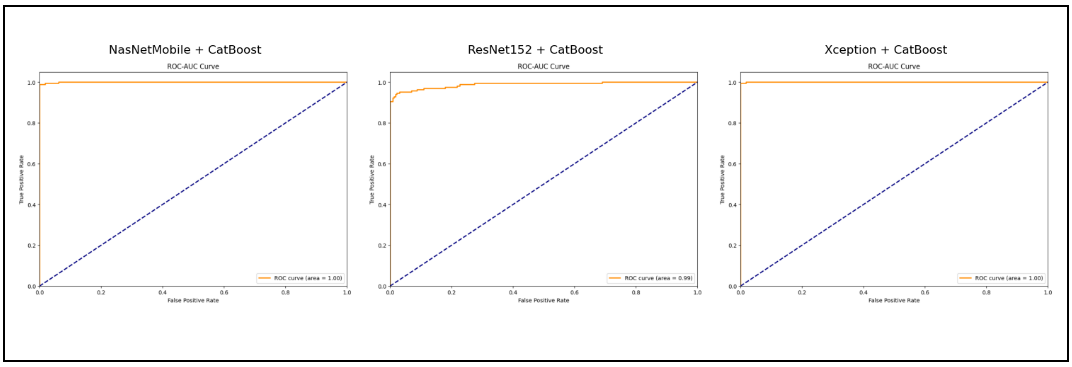


Figure 17: ROC-AUC Curve Analysis for Hybrid Model.

Figure 17 shows the ROC-AUC graph for Hybrid Models. All three perform well in distinguishing between classes with Xception + CatBoost and NasNetMobile + CatBoost showing a perfect score of 1.00. ResNet152 + CatBoost shows a close score of 0.99.

Table 6: Performance metrics for Hybrid CNN-Catboost

Model	Metric	Normal	PCOS	Accuracy
Xception + CatBoost	Precision	0.97	1.00	98%
	Recall	1.00	0.96	Specificity:100%
	F1-score	0.99	0.98	Sensitivity:96%
	Confusion Matrix	TN: 229, FN: 6	TP: 163, FP: 0	
ResNet152+ Catboost	Precision	0.91	1.00	94%
	Recall	1.00	0.86	Specificity:100%
	F1-score	0.95	0.93	Sensitivity:86%
	Confusion Matrix	TN: 229, FN: 23	TP: 146, FP: 0	
NasNetMobile + CatBoost	Precision	0.90	1.00	94%
	Recall	1.00	0.85	Specificity:100%
	F1-score	0.95	0.92	Sensitivity:85%
	Confusion Matrix	TN: 229, FN: 25	TP: 144, FP: 0	

Table 6 shows the evaluation matrix for Hybrid models built using CNN and Machine learning classifier CatBoost. The Xception + CatBoost model gives the best results with an accuracy of 98% giving perfect precision and recall which are reflected in the F1-scores of 99% for Normal and 98% for PCOS. The model has minimal misclassifications with only 6 false negatives for Normal and no false positives for PCOS. In contrast, RestNet152 + Catboost shows a slight dip in performance with an accuracy of 94% with 23 misclassifications for PCOS with a recall of 86%. NasNetMobile + CatBoost follows a similar pattern with an accuracy of 93% and 40 misclassifications for PCOS. Both ResNet152 and NasNetMobile show perfect precision for PCOS but differ in their ability to detect PCOS cases correctly.

6.6 Discussion

Table 7: Evaluation Metrics for Original Images

Model	Precision	Recall	F1-score	Accuracy
NasNetMobile	0.98	0.97	0.97	97%
ResNet152	0.98	0.97	0.97	97%
Xception	0.99	0.98	0.99	99%
Xception + CatBoost	0.99	0.99	0.99	99%
ResNet152 + CatBoost	0.98	0.97	0.97	97%
NasNetMobile + CatBoost	0.95	0.93	0.94	94%

Table 7 presents evaluation metrics on the classification results obtained by using the original data. In this research, the primary focus was on assessing the impact of SRGAN-generated data on PCOS detection using both CNN and Hybrid models. This approach was designed to understand if the addition of synthetic data can enhance the model’s performance in comparison with the performance of the original data. The evaluation metrics in Table 7 explains that with just original data models like NasNetMobile, ResNet152, Xception and the hybrid models combining these CNN models with CatBoost showed excellent performance achieving an accuracy of 99%. When integrating with SRGAN-generated data it was observed that although XceptionNet maintained a high performance of 99% for both standalone models, the performance for other models decreased a little to 96%. For hybrid models, the performance decreased to 94% indicating that CNN models like ResNet152 and NasNetMobile with CatBoost might have not been able to capture subtle details present in SRGAN-generated data. This suggests that while the synthetic data mostly matched the performance of original data, a slight decline was observed for certain models. This indicates that the impact of classification from SRGAN-generated data can slightly vary based on the model characteristics and architecture.

7 Conclusion and Future Work

In this research, attempts were made to leverage SRGAN-generated images to enhance model training but the synthetic images retained some blurriness, but if observed the original images were also a bit blurry and this can be due to the nature of the ultrasound images which are usually noisy and of complex nature. Based on these characteristics, SRGAN might not be the best choice for improving ultrasound images, but the study still explored this approach to evaluate its impact and effectiveness. The study only tested 4 variations for SRGAN due to computational constraints, suggesting that further exploration with more training and resources can potentially yield better results. In evaluating the effectiveness of SRGAN-generated data, CNN models including NasNetMobile, ResNet152, and Xception were tested alongside hybrid models combining these CNNs with catboost.

The models showed good accuracy even with the addition of SRGAN-generated data but there were minor differences in performance, with NasNetMobile and ResNet152 experiencing slight variations of a few percent in accuracy as compared to classification using original image. A conclusion can be made that if the images were poor quality, there would be a significant drop in performance compared to the original one, but the minimal difference suggests that the SRGAN-generated images in this research were of sufficient quality to be useful, but the performance can vary depending on the model architecture.

References

Ahmad, W., Ali, H., Shah, Z. and Azmat, S. (2022) 'A new generative adversarial network for medical images super resolution', *Scientific Reports*, 12(1), p. 9533. doi: 10.1038/s41598-022-13658-4.

Alamoudi, A., Khan, I.U., Aslam, N., Alqahtani, N., Alsalf, H.S., Al Dandan, O., Al Gadeeb, M. and Al Bahrani, R. (2023) 'A deep learning fusion approach to diagnosis the polycystic ovary syndrome (PCOS)', *Applied Computational Intelligence and Soft Computing*. doi: 10.1155/2023/9686697.

Chitra, P., Srilatha, K., Sumathi, M., Ishwarya, C. and Jagadesh, M. (2023) 'Automated detection of polycystic ovaries using pretrained deep learning models', in 2023 Annual International Conference on Emerging Research Areas: International Conference on Intelligent Systems (AICERA/ICIS). Kanjirapally, India, 16-18 November 2023, pp. 1-6. doi: 10.1109/AICERA/ICIS59538.2023.10420280.

Deshpande, S. S. and Wakankar, A. (2014) 'Automated detection of Polycystic Ovarian Syndrome using follicle recognition', in 2014 IEEE International Conference on Advanced Communications, Control and Computing Technologies. Ramanathapuram, India, 8-10 May 2014, pp. 1341-1346. doi: 10.1109/ICACCCT.2014.7019318.

Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J. and Greenspan, H. (2018) 'GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification', *Neurocomputing*, 321, pp. 321-331. doi: 10.1016/j.neucom.2018.09.013.

Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V. and Courville, A. (2017) 'Improved training of Wasserstein GANs', NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems. Long Beach, California, USA, 4-9 December 2017, pp. 5769 – 5779.

Hosain, A. K. M. 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', in 2022 International Conference on Engineering and Emerging Technologies (ICEET). Kuala Lumpar, Indonesia, 27-28 October 2022. doi: 10.1109/ICEET56468.2022.10007353.

Jia, P., Huang, Y., Cai, B. and Cai, D. (2019) 'Solar image restoration with the CycleGAN based on multi-fractal properties of texture features', The Astrophysical Journal Letters, 881(2), L30. doi: 10.3847/2041-8213/ab365f.

Kakoly, N.S., Earnest, A., Teede, H.J., Moran, L.J. and Joham, A.E. (2019) ‘The impact of obesity on the incidence of type 2 diabetes among women with polycystic ovary syndrome,’ *Diabetes Care*, 42(4), pp. 560-567. doi: 10.2337/dc18-1738.

Kumar, A. S., Annamalai, S., Kumaresan, M., Manikandan, P., Sekaran, R. and Pai, H. A. (2023) 'CNN-based analysis of ultrasound images for PCOS diagnosis', in 2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS). Tashkent, Uzbekistan, 1-3 November 2023, pp. 347-350. doi: 10.1109/ICTACS59847.2023.10390451.

Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z. and Shi, W. (2017) 'Photo-realistic single image super-resolution using a generative adversarial network', Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4681-4690.

Liang, J., Yang, X., Huang, Y., Li, H., He, S., Hu, X., Chen, Z., Xue, W., Cheng, J. and Ni, D. (2022) 'Sketch guided and progressive growing GAN for realistic and editable ultrasound image synthesis', Medical Image Analysis, 79, 102461. doi: 10.1016/j.media.2022.102461.

Lim, B., Son, S., Kim, H., Nah, S. and Lee, K. M. (2017) 'Enhanced deep residual networks for single image super-resolution', insum Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pp. 136-144.

Liu, S., Wang, Y., Yang, X., Lei, B., Liu, L., Li, S.X., Ni, D. and Wang, T. (2019) 'Deep learning in medical ultrasound analysis: A review', *Engineering*, 5(2), pp. 261-275. doi: 10.1016/j.eng.2018.11.020.

Manickam, S., Perumal, C., Sheela, S. and Ishwarya, C. (2023) 'Study and implementation of automated system for detection of PCOS from ultrasound scan images using artificial intelligence', doi: 10.1080/13682199.2023.2229016.

Maqsood, M. H., Mumtaz, R., Haq, I. U., Shafi, U., Zaidi, S. M. H. and Hafeez, M. (2021) 'Super resolution generative adversarial network (SRGANs) for wheat stripe rust classification', Sensors, 21(23), 7903. doi: 10.3390/s21237903.

Miyato, T., Kataoka, T., Koyama, M. and Yoshida, Y. (2018) 'Spectral normalization for generative adversarial networks', *arXiv*. doi: 10.48550/arXiv.1802.05957.

Moran, M.B.H., Faria, M.D.B., Giraldi, G.A., Bastos, L.F. and Conci, A. (2021) 'Using super-resolution generative adversarial network models and transfer learning to obtain high resolution digital periapical radiographs', *Computer Biology and Medicine*, 129, p. 104139. doi: 10.1016/j.compbiomed.2020.104139.

Nandhini, P. S., Srinath, P. and Veeramanikandan, P. (2022) 'Detection of glaucoma using convolutional neural network (CNN) with super resolution generative adversarial network (SRGAN)', in 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC). Trichy, India, 20-22 October 2022, pp. 1034-1040. doi: 10.1109/ICOSEC54921.2022.9951876.

Purnama, B., Wisesti, U. N., Adiwijaya, Nhita, F., Gayatri, A. and Mutiah, T. (2015) 'A classification of polycystic ovary syndrome based on follicle detection of ultrasound images', in 2015 3rd International Conference on Information and Communication Technology (ICoICT). Nusa Dua, Bali, Indonesia, 27-29 May 2015, pp. 396-401. doi: 10.1109/ICoICT.2015.7231458.

Ravishankar, T.N., Jadhav, H.M., Kumar, N.S., Ambala, S. and Pillai, M.N. (2023) 'A deep learning approach for ovarian cysts detection and classification (OCD-FCNN) using fuzzy convolutional neural network', *Measurement: Sensors*, 27, p. 100797. doi: 10.1016/j.measen.2023.100797.

Shashank, H. S., Acharya, A. and Sivaraman, E. (2023) 'Facial image super resolution and feature reconstruction using SRGANs with VGG-19-based adaptive loss function', in 2023 International Conference on Artificial Intelligence and Applications (ICAIA) Alliance Technology Conference (ATCON-1). Bangalore, India, 21-22 April 2023, pp. 1-6. doi: 10.1109/ICAIA57370.2023.10169373.

Srivastav, S., Guleria, K. and Sharma, S. (2024) 'A transfer learning-based fine tuned VGG16 model for PCOS classification', in 2024 2nd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT). Bengaluru, India, 4-6 January 2024, pp. 1074-1079. doi: 10.1109/IDCIoT59759.2024.10467747.

Subramani, S., Rarichan, A. and Chaithra, S. H. (2023) 'Utilizing deep learning techniques for detection of polycystic ovarian syndrome using ovarian ultrasound image', in 2023 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES). Chennai, India, 14-15 December 2023, pp. 1-7. doi: 10.1109/ICSES60034.2023.10465547.

Suha, S.A. and Islam, M.N. (2022) 'An extended machine learning technique for polycystic ovary syndrome detection using ovary ultrasound image', *Scientific Reports*, 12(1), p. 17123. doi: 10.1038/s41598-022-21724-0.

Takano, N. and Alaghband, G. (2019) 'SRGAN: Training dataset matters', *arXiv*, doi: 10.48550/arXiv.1903.09922.

Teede, H., Deeks, A. and Moran, L. (2010) ‘Polycystic ovary syndrome: A complex condition with psychological, reproductive and metabolic manifestations that impacts on health across the lifespan’, *BMC Medicine*, 8, 41. doi: 10.1186/1741-7015-8-41.

Wolf, W.M., Wattick, R.A., Kinkade, O.N. and Olfert, M.D. (2018) 'Geographical prevalence of polycystic ovary syndrome as determined by region and race/ethnicity', *International Journal of Environmental Research and Public Health*, 15(11), 2589. doi: 10.3390/ijerph15112589.