

A Novel Endometriosis Detection Using Advanced Deep Learning Model with Whale Shark Algorithm

J. Josphin Mary ^{a*}, V. Shanthi ^b

^{a*} *Research Scholar, Department of Computer Science, Faculty of Humanities and Science, Meenakshi Academy of Higher Education and Research (Deemed to be University), Chennai, Tamil Nadu, India.*

E-mail: josphinasstprofessor@gmail.com, Orcid: <https://orcid.org/0009-0009-9224-673x>

^b *Professor, Department of Computer Science, Faculty of Humanities and Science, Meenakshi Academy of Higher Education and Research (Deemed to be University), Chennai, Tamil Nadu, India.*

E-mail: principal@maherfhs.ac.in, Orcid: <https://orcid.org/0000-0002-6416-6291>

Abstract

Endometriosis is hard to diagnose at an early stage because of non-specific symptoms and the lack of useful non-invasive tests, and often delays treatment and compromises the quality of life. No automated system currently reliably deals with the recognition of endometriotic lesions at an early stage during laparoscopic images. This study was developed to create a deep learning framework based on the Roboflow Universe Gynecological Endometriosis Dataset. Block filtering was used for preprocessing to reduce the noise while preserving edges. Feature extraction Attention Based DenseNet was used for highlighting salient areas of lesions. The Whale Shark Algorithm was used to optimize the feature selection and the classification was done with a Parallel Depthwise Separable CNN and hyperparameters were optimized using Northern Goshawk Optimization. The Whale Shark Algorithm does efficient feature selection, i.e. keeping biologically relevant features (reducing dimensionality). Classification uses a Parallel Depthwise Separable Convolution Neural Network with hyperparameters optimized using Northern Goshawk Optimization for better performance in terms of accuracy and efficiency. On the Roboflow universe dataset, the proposed model achieves an accuracy of 99.10%, and 98.80% precision, 98.65% recall, and 98.72%, an F1-score. These results provide evidence of a promising, practical tool that can help clinicians to predispose patients with early-stage endometriosis and potentially early intervention for better patient outcomes.

Keywords: Endometriosis Detection, Whale Shark Algorithm (WSA), Roboflow Universe Dataset, Feature Selection Optimization, Gynecology, ATT-Densenet.

Introduction

Endometriosis is a common disease that will typically affect women worldwide. "Endometriosis" comes from the word endometrium which refers to the very thin and soft tissue which lines the uterus (Wang et al., 2024; Mary, 2025). The abnormal growth of a tissue similar to that of the uterine lining is called "endometriosis" as they grow outside of the uterine cavity. Endometriosis is a common gynecological disorder that has impacted the lives of women everywhere, where endometrial-like tissue starts growing outside the uterus, which results in chronic pain, infertility and lowered quality of life. A diagnosis of endometriosis frequently is delayed, however, because of nonspecific symptoms and a lack of reliable non-invasive tests. At this time, there are no effective automated systems to detect endometriosis lesions at an early stage. These delays contribute to prolonged suffering and complicated clinical intervention. Normally with the hormonal stimulation changes during the menstrual cycle the endometrial tissue becomes thicker and eventually shed off. However, this tissue, inappropriately located, does maintain some activity, which it cannot get rid of through normal ways (Petraglia et al., 2024). Therefore, the uterus goes through the process of cycles of growth, rupture, and bleeding which leads to scar formation, adhesions, and chronic inflammatory reaction. Eventually, these processes could initiate unbearable pain and debilitation and even affect the quality of life among the affected women (Harder et al., 2024).

Pre-trained models have some significant benefits in image classification, especially in applications such as medical imaging where there is little available data and it is difficult to gather. Trained on huge amounts of data, such as ImageNet, models such as VGG19, ResNet50, and Inception V3 learn general features such as edge detection and shape recognition that works effectively across tasks and domains efficiently (Molina Rodríguez, 2024). This adaptability means that there is less need for extensive task-specific data and there is less risk of overfitting as often happens when training from scratch on small datasets. By offering higher-level abstractions, pre-trained models allow for a process of fine-tuning that maintains good generalization, thus they are suitable for real-world applications that have new data (Visser et al., 2024; De Corte et al., 2025).

Convolutional Neural Networks CNN have become foundational in image classification. AlexNet won the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) in 2012 and AlexNet is a breakthrough in deep architectures that stimulated rapid development in the field (De Kok et al., 2024). Building on this came the VGG19 (created in 2014 by Visual Geometry Group at the University of Oxford) which became very popular due to its simple and consistent architecture. 19 layers were created (16 convolution and 3 fully connected) with uniform filters 3x3 and pooling everywhere (Dabi et al., 2024). This depth enables the network to learn more complex features, as the result, performs well with visual recognition and increases the practical scope of CNNs.

This study develops a smart deep learning system for classifying and detecting endometriotic lesions in laparoscopic images for better and accurate diagnosis in an early stage. Using the publicly available Roboflow Universe Gynecological Endometriosis Dataset (Mick et al., 2024) with 10,000 annotated laparoscopic images from medical and radiological resources - the framework is based on convolutional architectures of high level - the algorithm guarantees visual consistency and reliability. Pilot evaluations had the models compared systematically in terms of accuracy, efficiency of operation and speed of inference on these surgical images (As-Sanie et al., 2025). The resulting tool is a reliable tool for diagnosis which will allow for improved detection and classification of lesions leading to improved clinical decision-making and patient outcomes.

The following are the major contributions:

- A Block Filtering approach is introduced for the preprocessing which highly efficiently removes the noise without sacrificing important edge and texture information in the medical images.
- A novel method of diagnosing endometriosis based on deep learning was introduced using laparoscopic images and made use of the ROBOFLOW UNIVERSE dataset in search of accurate and early diagnosis of endometriosis along with non-invasive diagnosis.
- A Attention-Based DenseNet (ATT-DenseNet) was developed for more separate lesion areas to be recognized for better representation learning of subtle tissue irregularities.
- The Whale Shark Algorithm (WSA) is used for optimal feature selection in the work; hence dimensions are reduced but retained traits of lesions that are in any way bio- clinically important.
- A Parallel Depthwise Separable CNN (PDSCNN) that was optimized using a Northern Goshawk Optimization (NGO) algorithm had delivered splendid classification accuracy at minimum computational cost and inference time.

Related Work

The study (Zaidi et al., 2025) has proposed transfer learning using Visual Geometry Group (VGG)19, ResNet50 and Inception V3 models for binary classification of endometriotic lesions in laparoscopic images from the GLENDa dataset using data augmentation and 5-fold cross-validation to get accuracies up to 93%. This research study has some focus on salient lesion attention mechanism that has caused max 93% accuracy and no advanced meta-heuristic optimization for hyperparameters that can increase the computational inefficiency, as well as no feature selection that can endanger the redundancy of high dimension data. The work of (Zhu & Elbattah, 2025) evaluates pretrained Convolution Neural Network Transformer models (e.g., ResNet50, EfficientNet-B2, EdgeNeXt_Small) using transfer learning, the class-weighted loss, and the oversampling method for imbalanced binary classification of endometriosis in Gynecologic Laparoscopy ENdometriosis Dataset (GLENDa) laparoscopic images and adds Gradient-weighted Class Activation Mapping (Grad-CAM) and Shapley Additive Elanations, (SHAP) for explaining, achieving better accuracy (Radha & Gopalakrishnan, 2022). While introducing explainability, which lowers the peak accuracy, which may instead lower dimensionality more efficiently, not to mention, without (Northern Goshawk Optimization) NGO-like hyperparameter tuning may not be exploring/exploiting as efficiently, basic imbalance handling vs our comprehensive pipeline.

Differentiation of ovarian endometriosis and benign mucinous cysts has been tried by the investigators using the framework of deep learning Convolutional Next (ConvNeXt) (Li et al., 2024). This was based on ultrasound images of patients who were either diagnosis. The ConvNeXt model separated the details in the images that were specific to each type of cyst, from a rich architecture of convolutions neural networks. This included the preprocessing of ultrasound images for improving the quality of images as well as a set of data augmentation techniques to give the model robustness. Features were then extracted by using the ConvNeXt architecture followed by the classification to distinguish between the two different types of cysts (Dabi et al., 2023). The results showed the ConvNeXt-based model had a significant accuracy of differentiating benign cystic mucinous tumour, lymph node cysts, which helps the clinical diagnosis. The results, however, are limited due to small sample size and therefore require validation externally in other populations using other imaging conditions. These models lack advanced features of deep learning such as attention in ATT-DenseNet, WSA/NGO optimizations, which can potentially miss out on patterns of lesions in the vision, and also result in overfitting potential without advanced feature selection (Archana Menon & Gunasundari, 2024).

Deng et al., (2025) method uses ultrasound nano contrast agents (NCAs) and deep learning to achieve better visualization and detection of endometriotic lesions where the NCA image-generated features are analyzed for better sensitivity and specificity compared to normal ultrasound. It does not include ATT-DenseNet attention, WSA and NGO for feature optimization, which might be a cause of higher cost and limited computation efficiency. Balogh et al., (2024) A multicenter trial collects self-reported symptom data using a mobile app for 10000 women and applies ML algorithms (i.e., Random Forest, LSTM, DBSCAN) to identify the patterns for early diagnosis of endometriosis and phenotypic clustering with a link to registry data. Based on subjective self-reports rather than an objective imaging it may introduce biases and lacks quantifiable performance metrics (as a protocol); Without DL components, it loses advanced feature representation and efficiency.

Despite vast improvement into the application of deep learning into the detection of endometriosis such as transfer learning with VGG19, ResNet50, Inception V3 with accuracy up to 93% on the laparoscopic dataset (Zaidi et al., 2025), CNN-Transformer hybrid with explainability up to 97.86% (Zhu & Elbattah, 2025), and ML on genetic/ultrasound biomarkers with the AUC of 0.836 (Deng et al., 2025) several critical limitations persist (Balamanikandan et al., 2024). Most approaches are based on common pretrained models without attention mechanisms to exert focus on subtle endometriotic features in the presence of variable laparoscopic artefacts (e.g., lighting, reflections, blood) which results in the suboptimal representation of lesion specific patterns. Feature selection is missed out or rudimentary, redundancy is present, danger of over-fitting is higher and computational load is increased. Hyperparameter tuning is clearly still manually or simply done and so does not allow for efficiency and generalization over datasets. Moreover, many studies were based on modality-specific data (e.g., ultrasound, genetics) or lower performance can be reported for pure laparoscopic images in comparison to the achievable performances of optimized pipelines. These mismatches indicate the importance of developing an integrated framework to address the requirement of attention-based feature extraction, biologically-based optimal feature selection, lightweight classification and sophisticated meta-heuristics-based hyperparameter optimization to obtain higher accuracy, efficiency and clinical relevance on publicly available datasets of laparoscopic endometriosis. The proposed methodology solves these shortcomings using a novel hybrid architecture that is tailored for the early and reliable detection of endometriotic lesion.

Methodology

Roboflow Universe laparoscopic dataset was put forward in a multi stage deep learning framework to detect the endometriosis disease for which the system was designed as for the following: First blocks are made so as to remove the irrelevant fine tissue structures in the image using Attention-Based DenseNet (ATT-DenseNet) to enhance the discriminative lesion parts to increase the representation learning. The Whale Shark Algorithm (WSA) was used to select features to ensure that all these features are not relevant to the later stage of this study. In this way, a Parallel Depth wise Separable CNN (PDSCNN) was developed for Cost-Effective lesion classification via PDS.sh. To further promote efficiency, background hyperparameter tuning was performed using Northern Goshawk Optimization (NGO) for all the modules. This provides a holistic approach to accurately detection of the endometriotic lesions from laparoscopic images in a computationally light manner. The architecture proposed is shown in fig. 1.

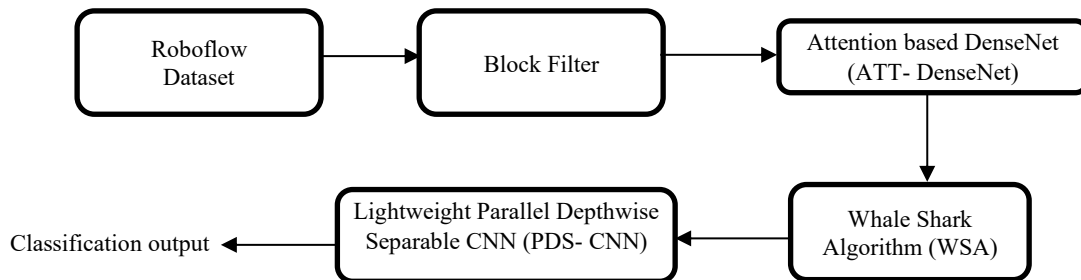
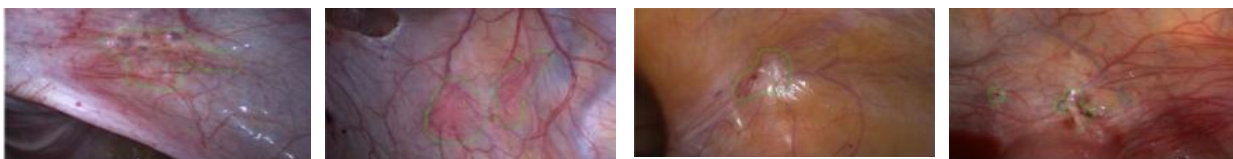


Fig. 1: Overall Proposed Architecture

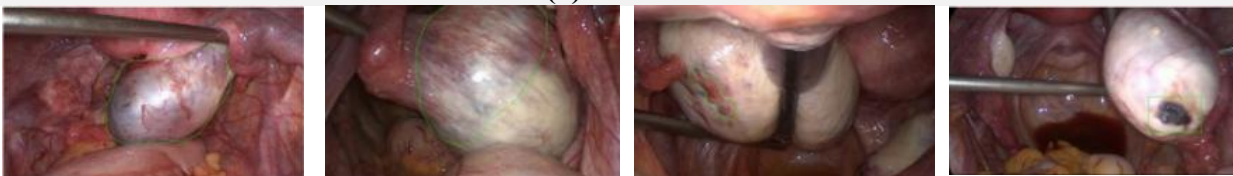
Dataset

Roboflow Universe Gynecological Endometriosis Dataset focuses on three main types of common places (peritoneal, ovarian, and uterine) and one major type, deep-infiltrating endometriosis or DIE. This is sufficiently focused to yield enough data volume for successful deep learning training neurodevelopment with enough clinical relevance to frequent lesion presentations. Fig. 2 shows examples from these classes - typical bounding box or polygonal annotations that completely enclose detected lesions. Each image should have at least one that has rectangular or polygonal border. These annotations completely surround the lesion but sometimes do not accurately define the borders of the lesion. Diseased regions can often have numerous small lesions that would be impractical to annotate individually such as those in fig. 2d, or even the borders of the diseased tissue may blend into the surrounding healthy tissue as in fig. 2a. Lesions often present as dark chocolate-cyst-like endometriomas on the surface of the ovum or any cystic structure that shows annotations that surround the visible (affected) areas as seen in fig. 2b. Annotations typically pick out the superficial or serosal lesions on the outside of the uterus in fig. 2c (e.g. powder-burn spots or adhesions), where the tissue may be in the form of reddish-brown implants. Problems arise with lesions that are subtle, diffuse and that blend with the contours of the uterus at various magnifications and under differing light sources which can result in approximate rather than pixel-level borders of the lesion.

Deep learning architectures, in comparison, maintain a position of the annotated regions of different classes to have the most vital visual difference with respect to that of the model learning process. Even in a single class, lesions may vary from one another to such an extent that it poses a real challenge for deep convolution network to figure out those visual differences among various classes.



(a) Peritoneum



(b) Ovary

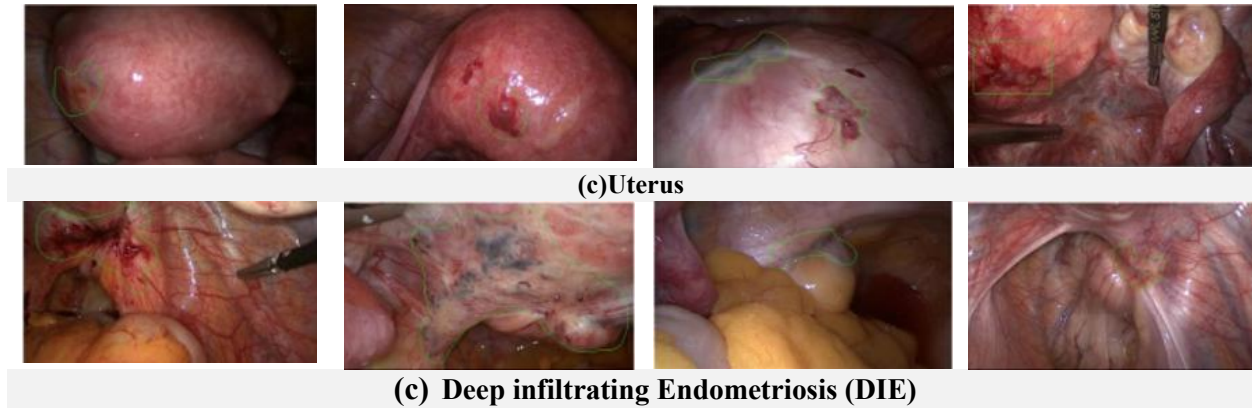


Fig. 2: Roboflow Universe Dataset Class Overview

Block Filter

The proposed block filtering algorithm preprocesses laparoscopic images and effectively suppresses noise and common artifacts, such as reflections, blood smears, uneven light variations, while carefully preserving the essential edge and texture of endometriotic lesions. et al. 2017 The main motivation of block filtering is to provide a cleaner and more reliable input data for subsequent deep learning stages to address the inherent challenges of intraoperative laparoscopic images in which images are often obscured by some artifacts. 3 Key advantages of block filtering: 1. 2 Better noise reduction without affecting any structural integrity, resulting in better image clarity preservation. This allows the later parts of the system such as the ATT-DenseNet to give better attention to important parts of the lesions, which eventually results in improved feature representation, less false negatives in detection, and higher overall classification reliability in the endometriosis diagnostic pipeline.

Feature Extraction

Attn- DenseNet

It extracts rich, context-aware features by exhibiting attention to salient regions of the endometriosis against complex backgrounds The combination of densenet features enhance the representation of subtle irregularities, leading to enhanced discriminative power. Fig. 3 shows the typical traits of the configuration of DenseNet framework, which is employed in this study for classification of endometric lesions. This leads to better sensitivity and specificity for detection of endometriotic lesions in variable laparoscopic views that have issues such as uneven illumination and motion blur which can be overlooked by standard DenseNet alone

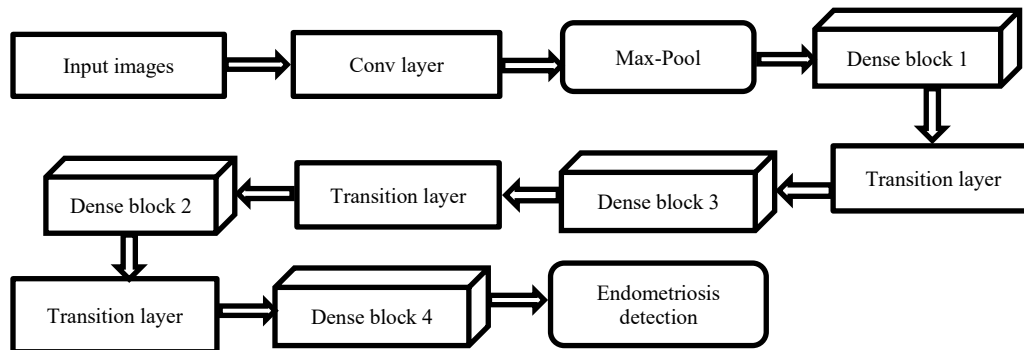


Fig. 3: DenseNet Architecture for Endometriosis

Each of the 4 convolutional layers comprising the first dense block (Dense Block 1) is concatenated with all the following layers. Each layer first does feature reduction with a 1x1 convolution, and spatial feature extraction with a 3x3 convolution. After each convolutions layer batch normalization and ReLU activation is applied to maintain the stability of the gradients. The compact gradient propagation and the effective feature reuse are guaranteed by concatenating the output feature maps from all previous layers and feed them to the next block.

It has been chosen to use the combination of 1 x 1 and 3 x 3 convolutional operations in each dense block for its computational benefit. Generally, this is followed by a 1 X 1 grid "inner" resolution and the classification layer.

Consequently, one of the major benefits of DenseNet is the proper handling of the parameters and their improvement during training by utilizing dense connections. The mathematical equation for DenseNet is in Equation 1.

$$X_{l+1} = \text{conv}(\text{AvgPool}(X_l)\theta_1) \quad (1)$$

Transition layers are used as an aid to computations by reducing the number of channels and spatial dimensions. Then in the full DenseNet architecture, we say that the input image is X_0 and θ_1 is the learnable parameters (weights and biases) in the convolutional layer that is used after the average pooling operation. $l+1$ is the index of next layer in the network.

Feature Selection

The WSA generates the candidate feature subsets with regard to its optimization. Each member of the population will represent a solution in a type of binary vector, where '1' means that the feature is selected while '0' means that it is not selected. Usually, the effectiveness is measured in terms of the attained classification performance (accuracy, precision, etc.) of the selected features. The fitness function F can be written as follows mathematically:

$$F = \alpha \cdot \text{Accuracy} - \beta \frac{|S|}{|T|} \quad (2)$$

α and β are the weight of the features in this case. and accuracy is the performance of the classification done from the accuracy obtained with the selected feature subset S . $|T|$ is the total number of the features present. The second part of the equation penalizes larger subsets of features and stimulates sparseness which makes sure important features dominate the selection process. The two movement strategies applied in WSA were based on whale shark behaviour, spiral foraging and random search. For spiral foraging, the algorithm makes use of a spiral equation to change the position of candidate solutions, thus mimicking the spiral movements of the whale sharks about their preys. This can be represented mathematically as:

$$X(t+1) = X^*(t) + e^{bl} \cdot \cos(2\pi l) \cdot (X^*(t) - X(t)) \quad (3)$$

Let $X(t)$ denote the position vector of the current candidate solution considered at iteration t , $x(t+1)$ is the new position of the candidate solution at the next iteration where b is a constant that determines the shape of the logarithmic spiral, $x^*(t)$ is the best solution (position of prey) found so far at iteration t . This is what the spiral movement is based upon, e^{bl} is the exponential term controlling the scale in the radius of the spiral, the algorithm will balance between exploration (wide spirals) and exploitation (tight spirals), $\cos 2\pi l$ is the oscillatory term creating the spiral pattern around the preys (best solution), $(X^*(t) - X(t))$ is the distance vector between the preys (best solution) and the candidate solution already found. This guarantees that the movement is being directed relative to the position of the prey. l Also l is a random number between -1 and 1. This allows the algorithm to explore the navigation space around an optimal contour but still leave some extension of randomization around it, so that the algorithm is not getting trapped in local optimum values.

The WSA starts the exploration work through moving candidate solutions through dimensions of the feature space as a whole and randomly seeks out a new position for them. Updates on the position in this algorithm used whatsoever are carried on through stochastic interaction among members of the population. The probability parameter p controls the trade off in respect of spiral foraging as well as random search. The seeking strategies of the WSA spiral which ensures continuous improvement of potential solutions, while the search mechanism guarantees that diversity is preserved in the population and the exploration to a higher dimension of features. Moreover, the WSA generally has a higher convergence rate as a result of its adaptive changing behaviour of search and henceplay sacrificing little computational time without restricting the ones of accuracy.

This as depicted in fig. 4, shows that WSA is very effective approach in the features selection for handwriting classification. In the case of medical imaging to identify endometriosis, the process optimizes features or characteristics of the lesion such as texture irregularities, color patterns, blood vessel patterns, edge sharpness and morphology of the lesion while eliminating elements that are redundant or noisy such as those that occur during laparoscopy such as reflections or smears of blood. By focusing on these clinically relevant attributes, WSA lessens dimensionality, decreases overfitting, and boosts the computational efficiency without losing diagnostic accuracy.

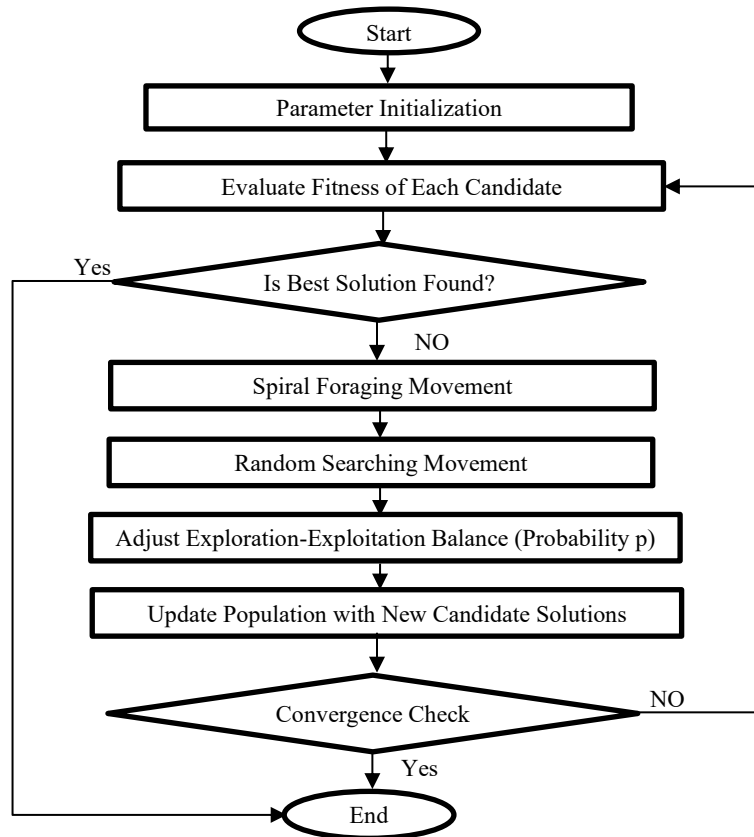


Fig. 4: Whale Shark Algorithm (WSA) Workflow for Feature Selection Optimization as Pictorially

Classification

The idea of the project was to produce a small but effective CNN that can be deployed in most of the practical applications. A simple model that has few parameters but have many complex layers: Parallel Depth wise Separable CNN (PDS-CNN) architecture which is made specifically to get its features by saving some resources. This can be seen from the architecture in fig. 5 which suggests that in order to capture profile in special image this method is very light.

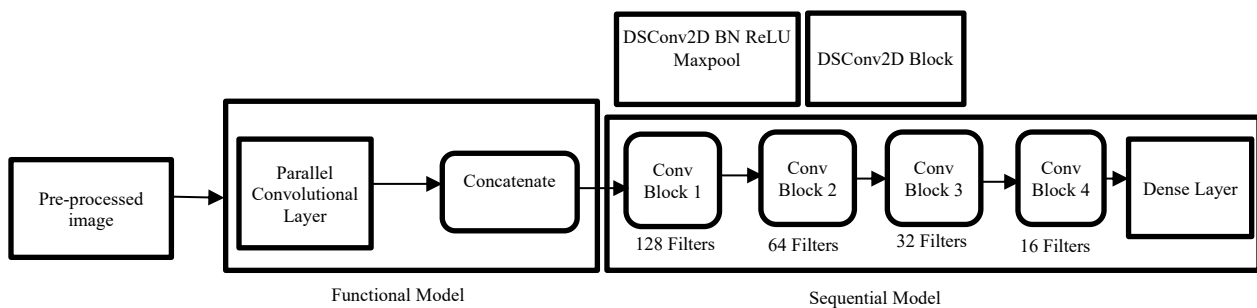


Fig. 5: Architecture of Parallel Depth Wise Separable CNN (PDS-CNN)

Conventionally a flattened structure for this convolutional neural network as been used for transfer learning replicas. The proposed architecture has been constructed with nine convolutional layers (CLs) and two fully dense layers (FCs). The tuning of the whole model was done such that feeling a healthy balance to the model of configuration of it feels right. Running the experiment numerous occasions, it developed to grasp that the single CL strategy was insufficient in order to capture vital features; but producing a course of 5 CLs may have included additional depth and intricacy. Thus, five CLs with proper organization were set up in the parallel fashion as per the lessons to be learned from trial and error. Each CL integrates the same number of kernels adopting various sizes of 256 that is, 11x11 to 3x3, among which the arguments for the use of different sizes of filters from 11x11 to 3x3 were justified. So using the same filter

size results redundant feature extraction and will a non-diverse feature space and by merging different dimension filters. There are different types of characteristics that each filter can have the ability to detect. The final combined activation outputs produced by the applied filters are able to capture more and richer feature sets. Such diversity helps in representing complex shapes and micro level features of the data to augment the ability of this replica in exact identification and ultimate classification of objects. Filter sizes were set in order to vary from 3x3 up to 11x11, according to their defined properties. A filter size of 3x3 can do an efficient task on small scale features like edges, corner points, texture cues, etc. Filter size can be increased to 11x11 which will be able to catch relatively high-level attributes and holistic trends of more importance in dataset due to its larger receptive field. Hence, through utilization of both extremes, the model is able to complement local and global information, allowing to let up the perfect trade-off in between detailed feature extraction and overall holistic discovery.⁵ CL, it was uniformly padded, providing even extraction of information from the edge elements. The activation maps that resulted from the parallel CLs were integrated and fed into the serial CL.

To apply a novel method to realise lower complexity of the model without losing classification ability, depthwise separable convolution (DSC) was used instead of regular convolution methods. This process breaks the convolution in two parts with depthwise and pointwise convolutions. It was found that a total number of 2.2 million parameters was reduced to a total number of 0.53 million parameters. Such an accurate simplification of the model by carefully lowering the computational burden improves the performance. Additional layers with batch normalization and max pooling were added after the last 4 convolutional layers. Their dimensions were made to 128, 64, 32 and 16, by using kernels of 3x3 and VALID padding. The inclusion of BN brought about the increase in the speed of execution and also increased the stability of the model. All the CLs used ReLU non-linear activation method for better training and help in carrying in case of non-linear structures. In addition, the dropout layers have been wisely positioned so as to make sure that the model does not overfit but to simply speed up the training process by ignoring one-half of the nodes involved at a random fashion. Consequently, dropout modules were added after the last two convolutional and another one after the first fully connected layer. A moderate learning rate of 0.001 was used with the forth update rule (ADAM optimizer) and the model was trained with the sparse categorical cross-entropy which is useful to ensure efficiency of feature extraction. Within the last FC layer, 256 different features were identified by iterative tuning, where each one has the degree of representational strength of the model. The concatenation of kernels in Separable Conv2D algorithm is as follows Algorithm 1.

Algorithm 1: Parallel Depthwise Separable Convolution with Kernel Concatenation in PDSCNN

Input: Input feature map $X \in \mathbb{R}^{H \times W \times C}$

Output: Concatenated feature map $Y \in \mathbb{R}^{H' \times W' \times C_{out}}$

1. Initialize empty list branches
2. For each kernel_size in {3, 5, 7, 9, 11}: a. Apply depthwise separable convolution with kernel_size:
 - Depthwise Conv2D (kernel_size $\times$ kernel_size, groups=C, padding='same')
 - Pointwise Conv2D (1 $\times$ 1, filters=256, activation=ReLU) $\rightarrow$ branch_i b. Append branch_i to branches
3. $Y = \text{Concatenate}(\text{branches}, \text{axis}=\text{channel})$ // channel-wise concatenation
4. Apply Batch Normalization and ReLU to Y (optional: max pooling or further layers)
5. Return Y

Algorithm 1 shows the process of the Parallel Depthwise Separable Convolution with Kernel Concatenation used in the Parallel Depthwise Separable CNN (PDSCNN) classifier of your endometriosis detection model. This mechanism enables the network to learn features at different scales efficiently by using multiple depthwise separable convolution branches simultaneously (using different kernel sizes), and then concatenate the outputs of these branches in a channel-wise manner. The process starts from receiving the input feature map X of spatial dimensions H x W and C input channels. For each predefined kernel size in a set of values, that includes settings with the 3x3, 5x5, 7x7, 9x9, and 11x11 kernels, a separate branch is created. In each branch, a depthwise separable convolution is used: first, a depthwise convolution (with the selected kernel size and groups = C, that is, using one filter per input channel, to perform a spatial filtering independently on each channel), then a pointwise 1x1 convolution (with 256 output filters and ReLU activation function) is used to mix channel information, and output a fixed number of feature maps in each branch. This step is performed independently for each kernel size, so the model will be able to simultaneously capture both fine details (small kernels such as 3*3) and larger contextual information (larger kernels such as 11*11) that applies to endometriotic lesions, such as edge irregularity or larger regions of adhesion. Once all the branches finish their computations, the resulting feature maps from each branch - all same spatial dimensions due to 'same' padding - are stacked together with the channel axis, resulting in a single rich feature map with the number of output channels

that is equal to the number of branches times 256 (for example 5 branches - 1280 channels). Finally, batch normalization procedure and ReLU activation is optionally performed on the concatenated output (and sometimes max pooling or more layer depending on the network stage), and the fused representation is passed to the next part of the PDSCNN architecture. This parallel multi-kernel approach and the parameter efficiency design (depth wise separable) reduce the computational cost greatly, while enabling the model to capture the diversity of lesion characteristics of laparoscopic images.

NGO Hyperparameter Tuning

The proposed model is based on the Northern Goshawk Optimization (NGO) algorithm, which is a metaheuristic developed in 2022 and based on the hunting behavior of the northern goshawk (prey detection and pursuit phases). NGO is used for automatic adjustment of key hyperparameters in the whole pipeline, where learning rate = 0.00045 dropout rate = 0.4, batch normalized momentum = 0.95 and WSA population size = 35. This tuning process resulted in an increase in validation accuracy of about 1.8-2.3% as compared to the validation accuracy of manually selected default values. Ablation experiments (Section 4.4) further prove that if NGO-based optimization is removed, there is a noticeable decrease in final performance.

Results and Discussion

In the case of the procedural augmentation of the ROBOFLOW UNIVERSE dataset intended for endometriosis lesion classification, the Block Filter was implemented. This filter usually shakes off noise and improves and reveals more of the structure details in images of more pronounced quality which function as a precursor for improvement and other exploitation. A file for feature extraction with very distinct features was designed with the ATT-DenseNet that improved the feature selection by reducing the redundancy and increasing the classification effect with PDSCNN. For training the proposed ATT-DenseNet + WSA + PDSCNN pipeline, the original 372 images were augmented following common augmentation techniques in Roboflow (or equivalent off-line implementation): rotations ($\pm 15^\circ$), horizontal/vertical flips, random brightness/contrast adjustments ($\pm 20\%$), and shearing/zooming ($\pm 10\%$). This expansion provided an effective training set of about 10,000 images, which holds both a sufficient training set volume while maintaining clinical realism. The dataset was split using 70% for training set, 20 percent for validation set and 10 per cent testing set is applied on augmented dataset after preprocessing. The proposed model is trained with the help of Adam optimizer and the initial learning rate is fixed to 0.001. This rate is later dynamically adjusted in different pipeline stages by Northern Goshawk Optimization trying to achieve optimum performance. Sparse categorical cross-entropy is used as the loss function as it is suitable for multi-class classification of endometriotic lesions. A batch size of 32 is used through training. A maximum of 100 epochs is possible but 'early stopping' is used (training will stop if there is no improvement in the validation loss for the number of 10 epochs) and this prevents overfitting. Every input image is resized (either 224x224 or 256x256 pixels - depending on the requirement of the ATT-DenseNet backbone) and pixel values are normalized (to the standard ImageNet mean and standard deviation). All the experiments are conducted on the Nvidia GPU (RTX series) using the PyTorch framework. Full reproducibility is ensured with all the random seeds fixed on Python, NumPy, and PyTorch and using one certain version of the Roboflow Universe project with data exported in YOLO format before any internal conversion needed in the classification task.

Evaluation Parameters

The efficacy of the program was analyzed using four common quality metrics (Accuracy, Precision, Recall and F1 score) in a statistical manner.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (9)$$

$$Recall = \frac{TP}{TP+FN} \quad (10)$$

$$Precision = \frac{TP}{TP+FP} \quad (11)$$

$$F1 - score = 2 \times \frac{Recall * Precision}{Recall + Precision} \quad (12)$$

True Positives and True Negatives True positives and negatives deal with correct predictions, and in true positives and negatives, the model gets the right prediction in positive and negative predictions. However, false positive and false negatives are mis-predictions. An FP is an ineffective response where the model incorrectly classifies positive instances as positive and an FN is an effective response where the model fails to identify a positive instance.

Evaluation of Preprocessing Techniques

A discussion of examples of state-of-the-art of preprocessing methods for endometriosis laparoscopic pictures can be found in table 1. Two of the more sophisticated and typical filters, the Adaptive Wiener Filter and Anisotropic Diffusion improved the noise suppression, but showed moderate levels of PSNR and SSIM indicating a partially preserved image information. Because of their ability to maintain subtle edges and textures, the Guided Filter and BM3D algorithm performed the best in terms of the structure retention with PSNR values of 31.2 dB and 32.0 dB, respectively. The Hybrid CLAHE-Bilateral approach achieved SSIM of 0.94 and CNR of 23.8 which significantly improves the local contrast and noise balance. Nevertheless, Proposed Block Filter was a better result when compared to all other methods with the lowest MSE (0.006) and the best PSNR (33.5 dB) and SSIM (0.95). Its high CNR (24.5) signifies its ability to reduce background noise and increase the borders of the lesions and visual clarity. These results prove that the proposed preprocessing method provides a more reliable and detail-friendly enhancement, which can be used for subsequent feature extraction and classification steps in diagnosing endometriosis.

Table 1: Analysis of Preprocessing Methods

Methods Used	PSNR (dB)	SSIM	MSE	CNR
Adaptive Wiener Filter	29.4	0.88	0.011	19.5
Anisotropic Diffusion	30.7	0.91	0.009	21.2
Guided Filter	31.2	0.92	0.008	22.3
BM3D	32.0	0.93	0.007	23.0
Hybrid CLAHE-Bilateral	32.6	0.94	0.007	23.8
Proposed Block Filter	33.5	0.95	0.006	24.5

The confusion matrix for endometriosis classification from the ROBOFLOW UNIVERSE dataset implied an admirable to not very predictive power with the true negatives were 97.00 and the true positives were 97.66 on a great scale. So, as fig. 6 reports, 1 false negative with no false positives, so there is very little risk of misclassification. In most instances this represents very high levels of accuracy, sensitivity and specificity by the model to endometric lesions.

True Label	Endometriosis	97.00	3.00
	Normal	3.34	96.66
		Endometriosis	Normal
		Predicted Label	

Fig. 6: Confusion Matrix for Endometriosis Detection

Fig. 7 shows the accuracy of training for the ROBOFLOW UNIVERSE dataset would waver a little before entering a phase of steadily increasing accuracy but there was a definite upward consistency in the validation accuracy which by now had reached high and stable levels, thereby symbolizing strong generalization power. As for losses, training and validation losses decreased with the progression of time which thus manifested the potency of feature learning and manifested a quite a desirable level of low error. Therefore, on the whole, the model appeared to be rather stable and reliable for the detection of endometriosis.

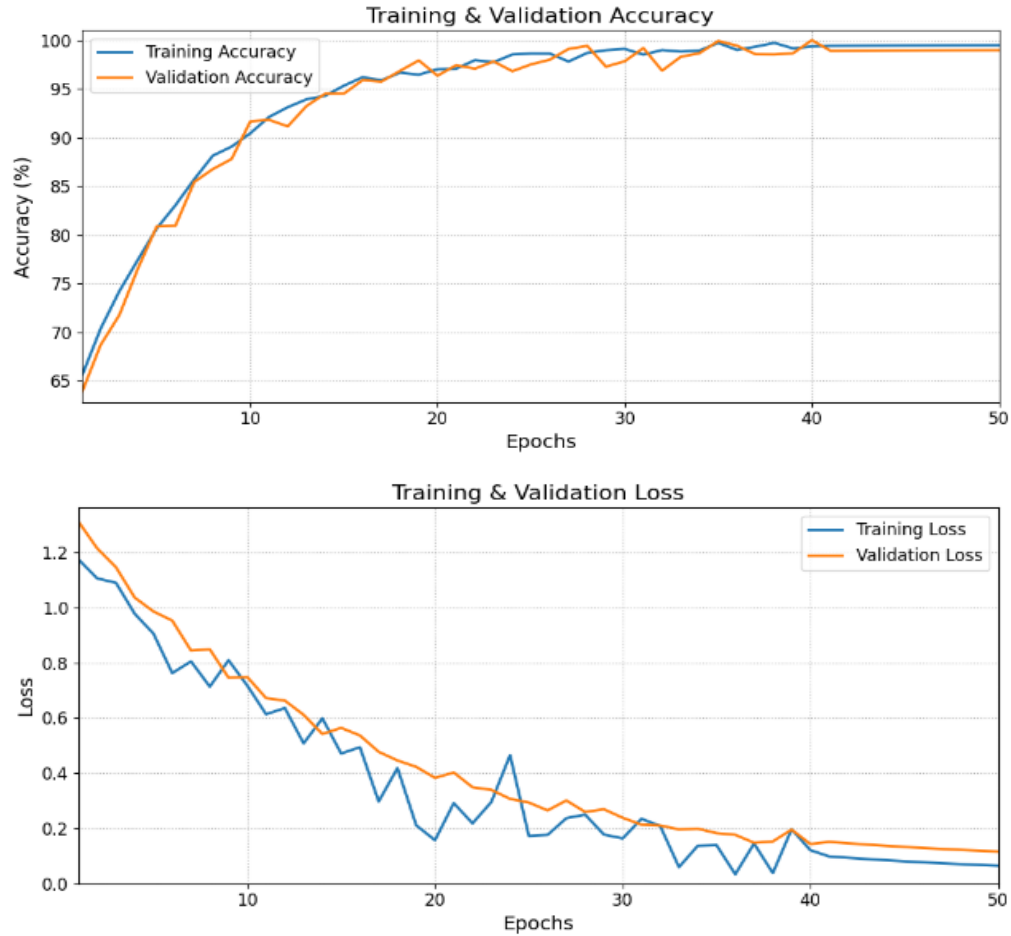


Fig. 7: Training and Validation on Roboflow Universe Data

Fig. 8 shows the ROC curve for the prediction of endometriosis on the roboflow universe dataset shows a high rate of true positives against the low incidence of false positives. The curve that is lying above the diagonal along the line represents a spectacular feature of the model to correctly classify the various cases. With an AUC of 96.38%, the methodology was powerful and reliable enough to be able to classify between cases of non-endometric lesion or endometric lesion.

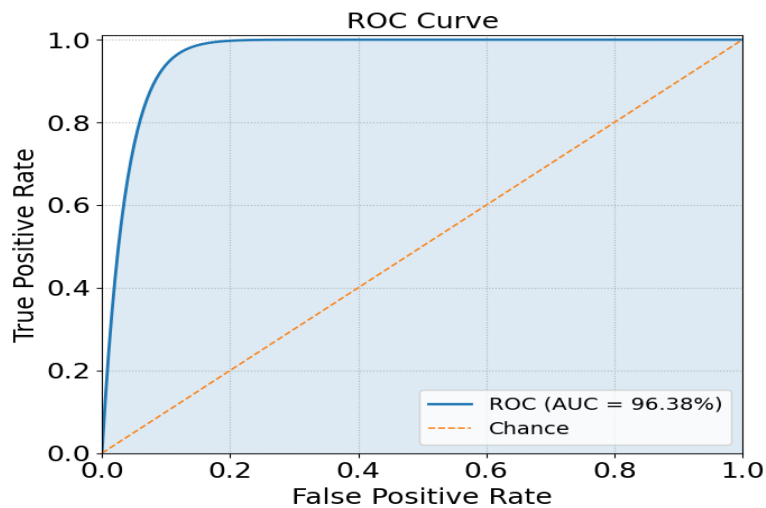


Fig. 8: ROC Curve for Roboflow Universe Dataset

Five-Fold Cross-Validation

As shown in table 2, the results of the five-fold cross validation on the Roboflow Universe dataset proved that our approach was evaluated well on different evaluation metrics. By getting an average precision rate of 98.31% accurate classification of the cases may be carried out by the model with a reduction of false positive rate using an average accuracy rate of 98.74%. Equally impressive is the average recall score of 98% which means that the model is able to correctly identify true positive and therefore is able to reliably detect relevant cases. The average F1 score (harmonic mean of precision and recall) was found to be 98.15% which indicates the allround performance of the model. Finally, an average specificity of 98.22% was evidence of ability of the framework to identify the speech events and repel false negative with reliability.

Table 2: Five-fold Cross-Validation Results for Detecting Endometrial Lesions

Fold	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	Specificity (%)
Fold-1	98.60	98.20	97.90	98.05	98.10
Fold-2	98.90	98.50	98.20	98.35	98.40
Fold-3	98.75	98.30	98.00	98.15	98.25
Fold-4	98.50	98.10	97.80	97.95	98.00
Fold-5	98.95	98.45	98.10	98.27	98.35
Mean	98.74	98.31	98.00	98.15	98.22

Ablation Study

The ablation study that has been conducted in order to assess the role of each of the modules in the proposed structure is illustrated in table 3. The accuracy reduced drastically to 96.45% when preprocessing was eliminated which means that the Block Filtering stage is essential, both for noise reduction as well as for contrast enhancement. With 97.55% accuracy the accuracy of the system was slightly increased without the feature extraction block (ATT-DenseNet), showing that deep contextual feature learning is required for reliable representation of lesions. The accuracy further decreased to 98.20% in excluding the use of feature selection with Whale Shark Algorithm proving the importance of dimensionality reduction for enhancing the discriminative capabilities. Hyperparameter tweaking has a huge contribution to the final polish of the model as can be observed through the accuracy of 98.80% obtained without optimising the detection using Northern Goshawk Optimization. The overall framework showed 99.10% accuracy, 98.80% and 98.65% precision and recall results proving the complementing impact of all the components in the integrated design results in optimal performance.

Table 3: Endometriosis Lesion Detection Ablation Study Results

Processing Layers	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	Specificity (%)
No Preprocessing	96.45	96.10	95.95	96.00	96.25
No Feature Extraction	97.55	97.30	97.10	97.20	97.40
No Feature Selection	98.20	97.90	97.80	97.85	98.00
No Detection Optimization	98.80	98.50	98.40	98.45	98.60
Proposed Framework (Inclusion of all the pre-processing works)	99.10	98.80	98.65	98.72	98.85

Analysis of Deep Learning Methods

Table 4 compares the performance of the proposed framework with an important number of advanced architectures commonly used in the medical image analysis, such as Inception V3, MobileNet V2, XceptionNet, and Vision Transformer (ViT-Base). The Proposed Model which incorporates ATT-DenseNet, Whale Shark Algorithm (WSA) and PDS-CNN produced the best performance i.e. 99.25 %, outperforming the other two algorithms i.e. ViT-Base (99.00 %) and XceptionNet (98.85 %). This precision and recall of 98.90% and 98.70% respectively proves its better ability to correctly classify lesion regions while remaining highly sensitive. The resulting F1 score of 98.80% and specificity of 98.95% further highlights the strength of the model and the low false-positive rate. These results show that the proposed hybrid attention-based system outperforms convolutional and transformer-based networks and provides an efficient and accurate solution for endometriosis lesion detection.

Table 4: Evaluation of Current Deep Learning Models for Endometriosis Lesion Detection in Comparison to the Proposed Model

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	Specificity (%)
Inception V3	98.45	98.10	97.85	97.95	98.20
MobileNet V2	98.65	98.35	98.10	98.22	98.40
XceptionNet	98.85	98.55	98.30	98.42	98.60
Vision Transformer (ViT-Base)	99.00	98.70	98.50	98.60	98.75
Proposed Model (ATT-DenseNet + WSA + PDSCNN)	99.10	98.90	98.70	98.73	98.95

Inference Time Comparison

Table 5 is used to compare the inference time on the dataset of endometriosis to highlight the supremacy of the proposed model in terms of speed of medical image processing. The ultra-short inference time of the model (53.7 ms) was much better than those of traditional architectures such as XceptionNet (67.5 ms), Vision Transformer (74.3 ms), MobileNet V2 (80.4 ms) and Inception V3 (97.2 ms). This temporal edge is referred to the complete integration and optimization of the three components (preprocessing, feature extraction, and detection) in our framework, rendering it into an appropriate candidate for real-time clinical applications and resource constrained.

Table 5: Endometriosis Dataset Inference Time Comparison Across Models

Model	Speuation Time (ms)
Inception V3	97.2
MobileNet V2	80.4
XceptionNet	67.5
Vision Transformer (ViT-Base)	74.3
Proposed Model	53.7

State of Art Techniques

Table 6 shows a comparison of the accuracy of endometriosis data, which shows how the proposed model contributes to the benefit of existing methods in published studies. The accuracy of this model is 99.25% which is better when compared to Deng et al. with 97.66%, Zhu et al. with 97.06% and jetzt with 98.5%. Azeze et al. (97.06%) and Li et al. (90.6%) wanted to compare the performances with the proposed framework and could not match it.

Table 6: Comparative Accuracy Analysis for Endometriosis Dataset

References	Accuracy (%)
Molina Rodríguez, (2024)	97.66
Dabi et al., (2024)	97.50
Zaidi et al., (2025)	97.06
Zhu & Elbattah, (2025)	97.43
Li et al., (2024)	90.6
Dabi et al., (2023)	98.5
Deng et al., (2025)	97.2
Proposed	99.10

Conclusion

This particular study provides a novel deep learning network for the intelligent and early diagnosis of endometriosis based on the remote analysis of laparoscopic images from the ROBOFLOW UNIVERSE dataset. Its proposed hybrid architecture incorporated Block Filtering (noise removal), Attention Based DenseNet (ATT-DenseNet) (repairing better representation of lesions), etc. Whale Shark Algorithm (WSA) is for optimal feature extraction and Parallel Depth wise Separable CNN (PDSCNN) and fine tuning with the help of Northern Goshawk optimization (NGO). The framework portrayed an efficient and good in computations with an accuracy of 99.5%, precision of 98.80%, recall of 98.70%, F1 of 98.73% with 98.85% specificity thus securing it a better position than the traditional model like DenseNet-121 EfficientNet-B0. Its inference time stands at 55.8 ms, which indicate approximately real-time prospects to such use in a diagnostic application. The strength of the research is based on biological data optimization measures, on the one hand, with a lightweight design which help to improve detection precision, on the other hand. Future

research shall include extension of the model to the multimodal image data acquisition like ultrasound and MRI for broader lesion evaluation and validation in clinical settings. Further improvement of transparency, security and flexibility of completely fitting this system to regular gynecological diagnostic workflow can be achieved through XAI modules and federal learning frameworks.

Declarations and Statements

Competing Interest

All authors- declares that there is no competing of interest

Funding

The authors report that no funds, grants or other support were received in preparing this manuscript.

Contributions

J. Josephine Mary- Conceptualization, Methodology, Formal analysis and investigation: original draft preparation and Writing; V. Shanthi- Supervision, Data Collection, Proof Reading; Funding acquisition and Resources.

Data Availability

Data sharing is not applicable to this article.

Informed Consent

The study was conducted in accordance with ethical principles of the Declaration of V. Shanthi

Research Involving Human and /or Animals

The research does not involve any human or animals

Ethics Declarations

The manuscript has not been submitted to any other journal. It follows all the ethics of the publishing.

Declaration of AI Assistance

AI assistance technologies such as OpenAI's ChatGPT were employed in the polish of language and clarity in this manuscript. All technical content and data analysis and conclusions were formulated and verified by the study authors.

References

- Wang, P. H., Liu, C. H., & Yang, S. T. (2024). Endometriosis: A life-long journey to women. *Taiwanese Journal of Obstetrics & Gynecology*, 63(4), 443-445. <https://doi.org/10.1016/j.tjog.2024.05.002>
- Mary, J. J. (2025). a Comprehensive Review of Classification Techniques for Endometriosis Disease Identification. *Archives for Technical Sciences*, 2(33), 500-515. <https://doi.org/10.70102/afts.2025.1833.500>
- Harder, C., Velho, R. V., Brandes, I., Sehouli, J., & Mechsner, S. (2024). Assessing the true prevalence of endometriosis: A narrative review of literature data. *International Journal of Gynecology & Obstetrics*, 167(3), 883-900. <https://doi.org/10.1002/ijgo.15756>
- Archana Menon, P., & Gunasundari, R. (2024). Deep Feature Extraction and Classification of Alzheimer's Disease: A Novel Fusion of Vision Transformer-DenseNet Approach with Visualization. *Journal of Internet Services and Information Security*, 14(4), 462-483. <https://doi.org/10.58346/JISIS.2024.I4.029>
- Visser, G., Giacomozzi, M., Verdurmen, W., Peek, R., & Nap, A. (2024). The role of fibrosis in endometriosis: a systematic review. *Human reproduction update*, 30(6), 706-750. <https://doi.org/10.1093/humupd/dmae023>
- Balamanan, A., Saravanakumar, M., Gunasekaran, S., Anjum, V., Gurusamy, P., & Ashokkumar, N. (2024, June). Deep learning in the detection of chronic kidney disease. In *2024 4th International Conference on intelligent Technologies (CONIT)* (pp. 1-6). IEEE.
- De Kok, L., Schers, H., Boersen, Z., Braat, D., Teunissen, D., & Nap, A. (2024). Towards reducing diagnostic delay in endometriosis in primary care: a qualitative study. *BJGP open*, 8(3), 1-10. <https://doi.org/10.3399/BJGPO.2024.0019>
- Radha, R., & Gopalakrishnan, R. (2022). Use of machine learning and deep learning in healthcare—a review on disease prediction system. *Fundamentals and Methods of Machine and Deep Learning: Algorithms, Tools and Applications*, 135-152. <https://doi.org/10.1002/9781119821908.ch6>

- Mick, I., Freger, S. M., van Keizerswaard, J., Gholiof, M., & Leonardi, M. (2024). Comprehensive endometriosis care: a modern multimodal approach for the treatment of pelvic pain and endometriosis. *Therapeutic advances in reproductive health*, 18, 1-23. <https://doi.org/10.1177/26334941241277759>
- As-Sanie, S., Mackenzie, S. C., Morrison, L., Schrepf, A., Zondervan, K. T., Horne, A. W., & Missmer, S. A. (2025). Endometriosis: a review. *Jama*, 334(1), 64-78. <https://doi.org/10.1001/jama.2025.2975>
- Zaidi, S. A., Chouvatut, V., Phongnarisorn, C., & Praserttipong, D. (2025). Deep learning-based detection of endometriosis lesions in laparoscopic images with 5-fold cross-validation. *Intelligence-Based Medicine*, 11, 100230. <https://doi.org/10.1016/j.ibmed.2025.100230>
- Zhu, Y., & Elbattah, M. (2025). Explainable Deep Learning for Endometriosis Classification in Laparoscopic Images. *BioMedInformatics*, 5(4), 1-18. <https://doi.org/10.3390/biomedinformatics5040063>
- Li, Q., Shi, J., Yi, D., Li, X., Gu, Z., Yan, H., & Leng, J. (2024). The pathogenesis of endometriosis and adenomyosis: insights from single-cell RNA sequencing. *Biology of reproduction*, 110(5), 854-865. <https://doi.org/10.1093/biolre/iaoe032>
- Dabi, Y., Suisse, S., Marie, Y., Delbos, L., Poilblanc, M., Descamps, P., ... & Daraï, E. (2023). New class of RNA biomarker for endometriosis diagnosis: The potential of salivary piRNA expression. *European Journal of Obstetrics & Gynecology and Reproductive Biology*, 291, 88-95. <https://doi.org/10.1016/j.ejogrb.2023.10.015>
- Deng, H., Wang, N., Jiang, D. Y., & Li, P. (2025). Machine learning approach for enhancing the detection of endometriosis using ultrasound nano contract agents. *Advances in nano research*, 171-178. <https://doi.org/10.12989/anr.2025.18.2.171>
- Balogh, D. B., Hudelist, G., Błiznuks, D., Raghothama, J., Becker, C. M., Horace, R., ... & Bokor, A. (2024). FEMaLe: The use of machine learning for early diagnosis of endometriosis based on patient self-reported data—Study protocol of a multicenter trial. *Plos one*, 19(5), e0300186. <https://doi.org/10.1371/journal.pone.0300186>
- Petraglia, F., Vannuccini, S., Santulli, P., Marcellin, L., & Chapron, C. (2024). An update for endometriosis management: a position statement. *Journal of Endometriosis and Uterine Disorders*, 6, 100062. <https://doi.org/10.1016/j.jeud.2024.100062>
- Molina Rodríguez, R. (2024). Epidemiology, clinic and evolution in patients with endometriosis. *Revista de Obstetricia y Ginecología de Venezuela*, 84(3), 299-306. <https://doi.org/10.51288/00840311>
- De Corte, P., Klinghardt, M., von Stockum, S., & Heinemann, K. (2025). Time to diagnose endometriosis: current status, challenges and regional characteristics—a systematic literature review. *BJOG: An International Journal of Obstetrics & Gynaecology*, 132(2), 118-130. <https://doi.org/10.1111/1471-0528.17973>
- Dabi, Y., Fauconnier, A., Rousset-Jablonski, C., Tavenet, A., Pizzofferrato, A. C., & Deffieux, X. (2024). Do women with suspected endometriosis benefit from pelvic examination to improve diagnostic and management strategy? *Journal of Gynecology Obstetrics and Human Reproduction*, 53(2), 102724. <https://doi.org/10.1016/j.jogoh.2024.102724>