Identification of Uterine Disorder using different Deep Learning Models

Baby Vijilin and Dr. Anitha V. S.

Govt.Engineering College Wayanad, Kerala Technological University, India

Abstract. Uterine disorders significantly impact women's reproductive health and overall well-being.The risk factors, early detection methods, and advancement in personalized treatment options, highlighting the importance of a multidisciplinary approach to manage these potentially life-threatening conditions. This paper details various research in diagnosing uterine disorders using a deep learning approach and a comparative study of various models in a deep learning approach in finding PCOS disorder from ultrasound images. Performance measures such as Accuracy, Precision, Recall, Confusion matrix, F1 Score, Training & Validation time and Execution time for prediction, are used to compare the results of different deep learning models.

Keywords: Uterine Disorders, Deep Learning, ultrasound, PCOS.

1 Introduction

Polycystic ovary syndrome (PCOS) is a hormonal disorder which is common in women of reproductive age. Due to PCOS, ovaries will be enlarged with small cysts. This hormonal imbalances likely influenced by genetic factors. Symptoms encompass irregular menstruation, elevated androgen levels, acne, and detectable ovarian cysts via ultrasound, potentially leading to fertility complications, insulin resistance, weight fluctuations, and increased risk of type II diabetes and cardiac related diseases. PCOS leads to the risk of obesity, infertility, and psychological issues like depression and anxiety. Treatment typically involves symptom management through lifestyle adjustments, hormone regulation medications, and fertility interventions.

This paper details various research undergone in diagnosing uterine disorders using deep learning approach and also comparative study of various models in deep learning approach in finding PCOS disorder from ultrasound images.Common benign uterine disorders are uterine fibroids, adenomyosis, and polycystic ovary syndromes. Malignant uterine disorders are endometrial and cervical cancers. In addition, attention is given to less common but clinically significant uterine disorders, including uterine anomalies, endometriosis, and uterine sarcomas. Diagnostic approaches, including imaging techniques and laboratory tests, are scrutinized for their accuracy and clinical utility in identifying these conditions. Magnetic resonance imaging (MRI), ultrasound, and hysteroscopy have emerged as pivotal tools in visualizing uterine anomalies, fibroids, and adenomyosis. The sensitivity and specificity of these modalities are critically evaluated, emphasizing their role in providing accurate and non-invasive diagnostic information.

Deep learning is a subset of machine learning. Deep learning models train artificial neural networks to learn patterns and predict from data. Deep learning has emerged as the major role in diagnostic field to find out various uterine disorders. In this section various diseases and deep learning approaches are discussed.

David C. Wyld et al. (Eds): AISO, SIP, SOFTFM, NLDM, CRBL, BIOM, COMIT– 2024
pp. 43-52, 2024. - CS & IT - CSCP 2024 DOI: 10.5121/csit.2024.141503

2 Literature Survey

20-50% of women around the world, have uterine leiomyomas or uterine fibroids which are the most common gynecological tumors. [1] Tumors in the female genital tract can be life threatening. [2]-[4]. Study of 122 female patients are included in [5]. Automatic segmentation can be done with the help of optimized U-net architecture. Presence of endometrial tissue outside the uterine cavity is known as Endometriosis, which is a disease of adolescents and reproductive-aged women. It is commonly associated with chronic pelvic pain and infertility. Endometriosis which can be symptomatic can be treated by hormone therapy and analgesics. Endometriosis often recurs, hence these treatments efficacy are limited. Identification of endometriosis using non-invasive in vitro tools minimize the diagnostic delay and hence reproductive health of patients can be improved[6].

Endometrial glands and stroma within the myometrium is known as Adenomyosis, which is a benign gynecological disease described by stroma within the myometrium, as well as reactive hyperplasia and hypertrophy of the muscular layer [6].

Detection of adenomyosis on uterine ultrasonographic images using Deep Learning approach[7] and detection of adenomyosis by the intermediate ultrasound skilled trainees are compared. Two architectures, ResNet and Vgg were considered in this work. Vgg13, Vgg19, ResNet 18 and ResNet 34 models were utilized in the work. Major two cancer types which threatens women's health are Uterine cervical and endometrial cancers .The deep learning (DL)-based research [8] tested the feasibility of classification of cervical and endometrial cancers and the site of origin of adenocarcinomas from whole slide images (WSIs) of tissue slides. [9] Performed clinical feasibility for the evaluation of fully automatic uterine segmentation on MRI by using images of major uterine disorders using U-net. T2-weighted MR images reveal the presence of uterine diseases, such as uterine leiomyoma, cervical cancer, and endometrial cancer etc. In this paper deals with automatic uterine segmentation of T2-weighted MR images.

Deep learning segmentation for ultrasonic images are referred in [10]. Uterine adenomyoma is identified using CNN, which is combined with the Deeplab network. The results of the FCN network are compared.The Deeplab network has obvious advantages as an image segmentation model of uterine adenomyoma.[7] Comparison of expertise of intermediate ultrasound skilled trainees and diagnostic performance of Deep Learning machine is stated. Cancer can be predicted using Deep learning and Bayesian optimization methods that have been remarkably successful. This research work discusses a risk backpropagation technique and the comparison of Bayesian optimized deep survival models and other state of the art machine learning methods.[11] These studies demonstrating improving accuracy and efficiency in medical practice with the application of deep learning techniques in uterine disorder diagnosis, classification, and prediction. Each of these papers has contributed significantly to the integration of deep learning methodologies in the domain of uterine disorders.

3 Deep Learning models in identifying PCOS

Deep learning models which are a subset of machine learning models. Deep learning models use neural networks to learn representations of data. These models have achieved significant success in image recognition. Here, we discuss some well-known deep learning architectures: VGG, ResNet, MobileNet, EfficientNet, Inception and EfficientNetB0 and they are applied to PCOS dataset from Kaggle to predict PCOS.

3.1 VGG

VGG (Visual Geometry Group) It is a convolutional neural network (CNN) architecture can be utilized for image classification.VGG has a simple architecture and its convolutional filters are having small size of 3x3 throughout the network, facilitating a large receptive field while keeping parameters manageable. VGG16 is using 16 layers and VGG19 is using 19 layers. The depth of these networks help to identify the complex structures and which lead to high performance on image classification tasks and various image vision tasks. VGG16 consists of 13 convolutional layers and 3 fully connected layers, while VGG19 includes additional layers. VGG is used for transfer learning to save computational resources and time.

VGG networks are computationally expensive due to their deep architecture and high memory consumption.The large size of VGG models can make them impractical for deployment in scenarios where storage space is limited, such as mobile or embedded devices.

Unlike more recent architectures like ResNet, VGG networks do not include residual connections, which help mitigate the vanishing gradient problem and improve the training of very deep networks. This lack of residual connections can limit the effectiveness of VGG networks in training very deep models .

3.2 ResNet

ResNet (Residual Network) introduces residual learning to enable the training of deep networks effectively. ResNet's hallmark is residual blocks, which incorporate skip connections or shortcuts, bypassing one or more layers to mitigate the vanishing gradient problem. Different variants are available for ReSNet such as ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152. It vary in layer depth to accommodate different computational resources and task complexities. Notably, ResNet50V2 utilizes residual learning and skip connections, comprising multiple residual blocks with convolutional layers. Features extracted from ResNet50V2 capture hierarchical visual information during training. These features serve as rich representations of the input image and can be used as input to subsequent layers or models for various computer vision tasks. High accuracy, scalability are some of the merits of the ResNet.Computational complexity and architectural complexity are some of the demerits of the ResNet. There is an increased risk of overfitting, especially when training on smaller datasets using very deep networks. This can be mitigated by Regularization techniques and data augmentation methods.

3.3 MobileNet

MobileNet, devised by Howard et al., targets efficient deployment on resource-limited devices. While comparing with traditional convolutional network, MobileNet uses depthwise separable convolutions, which help to reduce the number of parameters and computational cost compared to traditional convolutional networks.The lightweight architecture of MobileNet ensures faster inference times, thus it is appropriate for real-time applications on mobile and embedded devices. MobileNet introduces a width multiplier (α) and a resolution multiplier (ρ) that allow for trade-offs between latency, computational cost, and accuracy. This scalability makes it adaptable to various resource constraints and use cases. But one of the demerit of MobileNet is accuracy can be lower compared to more complex and larger models like ResNet or Inception, especially for more complex tasks or datasets.

3.4 Inception

Multi-scale features are captured effectively by Inception or GoogleMNet. These modules employ convolutional filters of various sizes (1x1, 3x3, 5x5) concurrently, enhancing feature representation. Additionally, 1x1 convolutions reduce input channels, managing computational complexity. Parallel operations process data through filters of different sizes concurrently, with outputs concatenated for comprehensive feature extraction. Inception models introduce auxiliary classifiers to combat the vanishing gradient problem, enhancing training convergence. Versions like InceptionV1 to InceptionV4 improve computational efficiency, accuracy, and training speed. The architecture is scalable, with subsequent versions (Inception-v2, Inception-v3, Inception-v4, and Inception-ResNet) incorporating various improvements and optimizations to enhance performance and efficiency. Inception models are versatile and can be adapted for various computer vision tasks, including image classification, object detection, and feature extraction in more complex systems. Inception models are versatile and can be adapted for various computer vision tasks, including image classification, object detection, and feature extraction in more complex systems. The depth and number of parameters in Inception models can lead to high memory consumption.Its complexity and resource requirements may challenge implementation and training.

3.5 Xception

Xception, an evolution from Inception, utilizes extreme versions of inception modules and depth-wise separable convolutions for feature extraction,separable convolutions, decoupling spatial and channel-wise filtering. This approach efficiently captures mid-level features, reduces the number of parameters, maintaining a balance between accuracy and computational cost leading to better performance. Xception includes global pooling layers for aggregating features across spatial locations and bottleneck blocks for parameter reduction while preserving representational capacity. These features make Xception suitable for various computer vision tasks, offering a different approach to feature extraction compared to Inception. Xception is well-suited for transfer learning, where pre-trained models on large datasets like ImageNet can be fine-tuned for specific tasks with smaller datasets, achieving good performance with less training data. The project may pose a risk of overfitting, especially when trained on small datasets and it requires careful tuning of hyperparameters to achieve optimal performance, which can be time-consuming and resource-intensive.

3.6 DenseNet

Densely Connected Convolutional Networks (DenseNet) represent a significant advancement in deep learning, particularly in the field of computer vision. Introduced by Huang et al. in 2017, DenseNets distinguish themselves through their unique architecture, where each layer is directly connected to every other layer in a feed-forward fashion. This dense connectivity pattern ensures maximal information flow between layers, facilitating feature reuse and mitigating the vanishing gradient problem, which is prevalent in very deep networks. By concatenating feature maps from all preceding layers, DenseNets enhance gradient propagation and reduce the need for redundant feature learning, leading to more efficient parameter usage compared to traditional CNNs. Additionally, the incorporation of transition layers between dense blocks helps control model complexity and spatial dimensions. These characteristics make DenseNets particularly effective for image classification, segmentation, and object detection tasks, demonstrating superior performance and efficiency in numerous benchmark datasets.DenseNet offers several advantages, including improved gradient flow, efficient parameter use, and rich feature representations. But it also faces challenges related to high memory consumption, increased computational complexity, and slower training times. Despite these challenges, DenseNet remains a powerful and effective architecture in the field of computer vision.

3.7 EfficientNetB0

EfficientNetB0 models are commonly used for transfer learning with smaller datasets. It has a highly efficient and scalable architecture with strong performance across various tasks. EfficientNetB0 achieves high accuracy on benchmark datasets like ImageNet, often outperforming larger models while using fewer parameters and computational resources.It requires fewer FLOPs (floating point operations), making it more suitable for deployment in resource-constrained environments such as mobile devices. It also presents challenges in terms of training complexity, memory consumption, and the need for extensive fine-tuning. Despite these challenges, EfficientNetB0 remains a powerful and versatile choice for modern computer vision applications.

4 Methodology and DataSet

4.1 Methodology

Our Aim is to identify the PCOS disease from the dataset. For this work we used the dataset which are available in Kaggle dataset. Several preprocessing , augmentation parameters and general augmentation techniques are applied to images for the standardization and normalization.

a) Images rescaled the pixel values of the images from the range [0, 255] to [0, 1]. It is a common practice in image preprocessing to normalize the pixel values to make the training process more stable and faster.

b) 20% of the data will be used for validation purposes, and 80% for training.

c)Images to be randomly rotated by up to 20 degrees. This helps the model to be invariant to small rotations and improves its generalization.

d)The images to be randomly shifted horizontally (left or right) by up to 20% of the width of the image. This helps the model to handle horizontal translations in the images.

e)The images to be randomly shifted vertically (up or down) by up to 20% of the height of the image. This helps the model to handle vertical translations in the images.

f)The images to be randomly sheared. Shearing is a type of transformation that tilts the image, which can help the model be more robust to distortions.

g)The images to be randomly zoomed in or out by up to 20%. This augmentation helps the model to be invariant to changes in scale.

h)The images to be randomly flipped horizontally. This is useful for images where the horizontal orientation is not important, such as most natural images.

i)The images to be randomly flipped vertically. This is less commonly used than horizontal flipping, as vertical orientation can be important for some types of images, but it can still be useful in some cases.

Above preprocessing methods are performed to enhance the robustness and generalization of the model by making it invariant to various image transformations and by providing a separate validation set for unbiased performance evaluation.

Different deep learning models are applied to this dataset and performance measures are calculated.

4.2 Dataset

Medical image datasets for uterine-related diseases are important for developing and training deep learning models in medical diagnosis. Kaggle datasets encompass a broad range of topics and come from diverse sources. Kaggle dataset is utilised for this work. The dataset is classified into two classes as “infected” and “notinfected” under PCOS disease category.

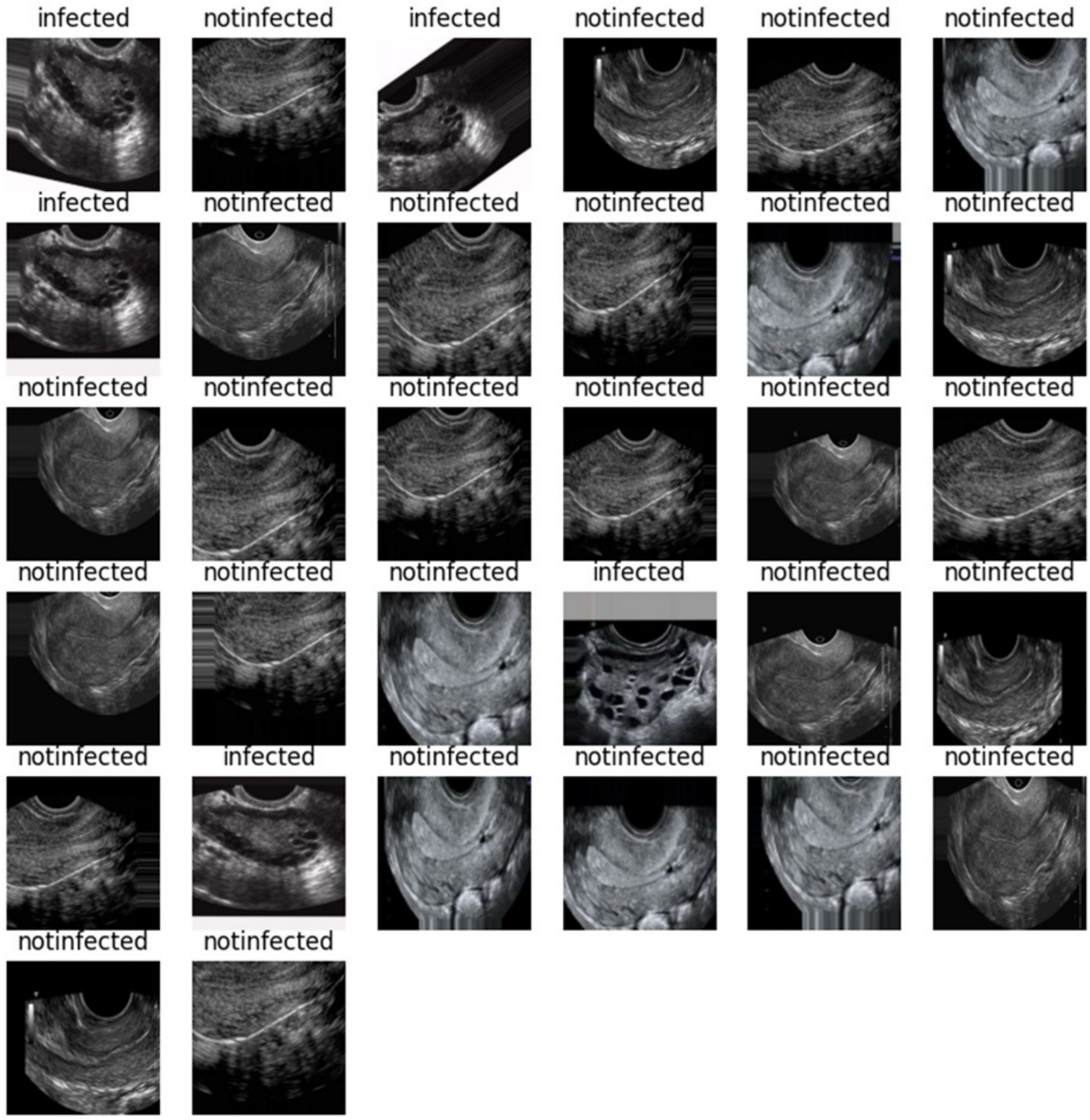

Fig. 1. PCOS Dataset.

5 Results and Discussions

The different deep learning techniques trained and validated using Kaggle dataset.(Figure 1) Various models in convolutional neural networks are applied into that dataset. Models applied in this work are VGG, Resnet50V2, Resnet50, MobileNet, InceptionV3, DenseNet,

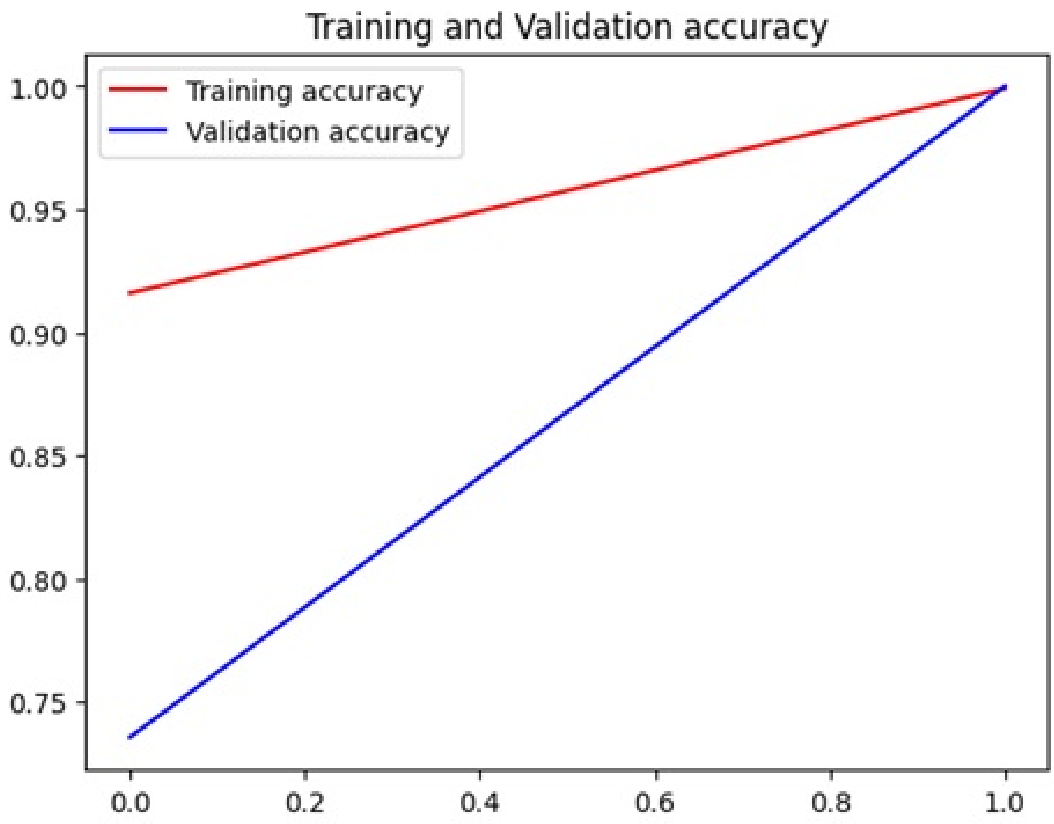

Fig. 2. MobileNet

Xception and EfficientNetB0. Performance measures used in this work are precision, recall, accuracy, F1Score, Confusion matrix and execution time. Performance measures of various models are listed in the Table 2.

5.1 Performance Measures

Following performance measures are used in result analysis.
1. Accuracy: It is the ratio of correct predictions to the total instances. It checks the correctness of the model. Formula:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

TP: True Positives TN: True Negatives FP: False Positives FN: False Negative 2. Precision: Precision is the ratio of correct predictions to the total predicted positives. It conveys whether the number of predicted positives are really positive.

$$\text{Precision} = \frac{TP}{TP + FP}$$

3. Recall (Sensitivity or True Positive Rate): Recall is the ratio of actual positives to all the observations in the actual class. It indicates how many of the actual positives were identified correctly.

$$\text{Recall} = \frac{TP}{TP + FN}$$

4. F1 Score: The F1 Score is the harmonic mean of Precision and Recall. It gives a balance between Precision and Recall. Model's accuracy on dataset can be measured using F1 score.

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{1}$$

5. Confusion Matrix: Performance of a model can be evaluated using confusion matrix table.It provides a view about the correct and incorrect classifications.

Actual / Predicted	**Positive (P)**	**Negative (N)**
Positive (P)	True Positive (TP)	False Negative (FN)
Negative (N)	False Positive (FP)	True Negative (TN)

Table 1. Confusion Matrix

True Positive (TP): The model correctly predicts the positive class.

True Negative (TN): The model correctly predicts the negative class.

False Positive (FP): The model incorrectly predicts the positive class (Type I error).

False Negative (FN): The model incorrectly predicts the negative class (Type II error).

6. Execution Time: It is the total time taken by a system or a process to complete a specific task. In the context of machine learning models, it refers to the time taken to train the model or to make predictions. Measurement is measured in seconds, milliseconds, or microseconds, depending on the granularity required.It can be calculated using various tools and programming constructs, such as:

6 Conclusion

The PCOS dataset correctly classified into two classes as "infected" and "notinfected" by all the models and the accuracy , precision and recall are ideal in all cases. Only the execution time found different for different models. From these models MobileNet is found better performance than other models while considering the execution time. While comparing the MobileNet model with least performing model , Mobilenet is having 300% better performance than the least performing model. Future Work can be made, improving the execution time with 100% accuracy of the predictions based on the performance measures.

Table 2. Comparative study of Performance measures of various model

Sl. No.	Model	Precision	Recall	Accuracy	F1 Score	Confusion Matrix	Building model training and validation time required for the dataset	Execution time for prediction
1	VGG	1	1	1	1	[[1 0] [0 1]]	9804.36s	5.942s
2	Customized VGG	1	1	1	1	[[1 0] [0 1]]	2431.879s	5.371s
3	Resnet50	1	1	1	1	[[1 0] [0 1]]	3473.97s	4.815s
4	Customized Resnet50	1	1	1	1	[[1 0] [0 1]]	1280.929s	5.037s
5	Resnet50V2	1	1	1	1	[[1 0] [0 1]]	3023.529s	5.681s
6	Customized Resnet50V2	1	1	1	1	[[1 0] [0 1]]	951.212s	4.944s
7	MobileNet	1	1	1	1	[[1 0] [0 1]]	1151.52s	3.138s
8	Customized MobileNet	1	1	1	1	[[1 0] [0 1]]	863.009s	4.47s
9	InceptionV3	1	1	1	1	[[1 0] [0 1]]	25.86s	4.872s
10	Customized InceptionV3	1	1	1	1	[[1 0] [0 1]]	963.63s	10.881s
11	DenseNet	1	1	1	1	[[1 0] [0 1]]	4625.418s	12.541s
12	Customized DenseNet	1	1	1	1	[[1 0] [0 1]]	1096.425s	8.556s
13	Xception	1	1	1	1	[[1 0] [0 1]]	8006.876s	4.362s
14	Customized Xception	1	1	1	1	[[1 0] [0 1]]	1169.022s	5.377s
15	EfficientNetB0	1	1	1	1	[[1 0] [0 1]]	1964.658s	3.493s
16	Customized EfficientNetB0	1	1	1	1	[[1 0] [0 1]]	1023.97s	8.479s

References

1. Woźniak A, Woźniak S. , "Ultrasonography of uterine leiomyomas," Prz Menopauzalny. 2017 Dec;16(4):113-117. doi: 10.5114/pm.2017.72754. Epub 2017 Dec 30. PMID: 29483851; PMCID: PMC5824679.
2. N. Hassani, Ultrasonic evaluation of uterine fibroids -The Sonic Shadow Sign, JOURNAL OF THE NATIONAL MEDICAL ASSOCIATION, Vol. 67, No. 4,307-311,1975.
3. K. Walker, K Jayaprakasan, N.J Raine-Fenning ,Ultrasound in benign gynaecology Obstetrics, Gynaecology & Reproductive Medicine, voI.17., no.2, pp.33- 44,2007.
4. G. Siskin,New Treatments for Uterine Fibroids, Techniques in Vascular and lnterventional Radiology, voI.9., no. l,pp. 12-18,2006.
5. KurataY. et al.Automatic segmentation of the uterus on MRI using a convolutional neural network,Comput. Biol. Med.(2019).
6. Cunningham R.K., Horrow M.M., Smith R.J., Springer J. Adenomyosis: A Sonographic Diagnosis RadioGraphics 2018;38:1576–1589. doi: 10.1148/rg.2018180080.

7. Raimondo D, Raffone A, Aru AC, Giorgi M, Giaquinto I, Spagnolo E, Travaglino A, Galatolo FA, Cimino MGCA, Lenzi J, Centini G, Lazzeri L, Mollo A, Seracchioli R, Casadio P. Application of Deep Learning Model in the Sonographic Diagnosis of Uterine Adenomyosis. Int J Environ Res Public Health. 2023 Jan 18;20(3):1724. doi: 10.3390/ijerph20031724. PMID: 36767092; PMCID: PMC9914280.
8. Song, J.; Im, S.; Lee, S.H.; Jang, H.-J. Deep Learning-Based Classification of Uterine Cervical and Endometrial Cancer Subtypes from Whole-Slide Histopathology Images. Diagnostics 2022, 12, 2623. https://doi.org/10.3390/diagnostics12112623.
9. Parasar P, Ozcan P, Terry KL. Endometriosis: Epidemiology, Diagnosis and Clinical Management. Curr Obstet Gynecol Rep. 2017 Mar;6(1):34-41. doi: 10.1007/s13669-017-0187-1. Epub 2017 Jan 27. PMID: 29276652; PMCID: PMC5737931.
10. Yi Jin, Wendi Huang, Qinghong Qu, "Analysis of Convolutional Neural Network Segmentation Algorithm of Adenomyoma", Computational Intelligence and Neuroscience, vol. 2022, Article ID 1629443, 8 pages, 2022. https://doi.org/10.1155/2022/1629443
11. Yousefi, S., Amrollahi, F., Amgad, M. et al. Predicting clinical outcomes from large scale cancer genomic profiles with deep survival models. Sci Rep 7, 11707 2017), https://doi.org/10.1038/s41598-017-11817-6.

Authors

Baby Vijilin Currently, she is pursuing her PhD in Computer Science from the Kerala Technological University.

Dr.Anitha V S received PhD from NIT Calicut. Her research interests include Networks, Image Processing and Deep Learning.