RESEARCH PAPER

A Novel DyDimNet-Based Dimension-Aware Deep Learning Framework for Automated Ovarian Tumor Classification Using Ultrasound Images

Aarti Parekh[1]**, Dr. Madhavi Desai**[2]
[1]Research Scholar, Gujarat Technological University, Ahmedabad, Gujarat, India.
Email: parekhaarti155@gmail.com
[2]Professor, CSE Department, R.N.G Patel Institute of Engineering, Tajpor, Bardoli, Gujarat, India.
Email: desaimadhavi30@gmail.com

ABSTRACT
Accurate identification of ovarian tumors is crucial in increasing treatment options, reducing mortality, and facilitating successful treatment. Due to their non-invasive nature and low cost, ultrasound imaging is extensively used in the identification of ovarian tumors. Unfortunately, many challenges in diagnostics such as speckle noise, variations in anatomy, class imbalance, and subtle differences between classes can significantly negatively impact diagnostic performance. Most current deep learning algorithms use uniform convolutions and generalized augmentations that limit the ability to utilize dimension-specific spatial features found in ovarian ultrasound images. To overcome these limitations, this study proposes the DyDimNet framework with Dimension-Aware Adaptive Augmentation (DAAA) for automated classification of ovarian tumors. DyDimNet improves discriminative feature learning through the dynamic analysis of spatial dependency across the height, width and depth dimensions. The DAAA approach applies augmentation based on physical dimensions in order to preserve physical fidelity and improve the representation of the minor class. The framework was evaluated using experimental tests on the MMOTU ovarian ultrasound dataset for both binary (96.8% accuracy) and multi-class (92.6%) classification tasks in an efficient computational manner and real-time feasibility for the respective data representation.

Keywords: Ovarian Tumor Classification, Ultrasound Imaging, Deep Learning, DyDimNet, Dimension-Aware Adaptive Augmentation (DAAA), EfficientNetV2, Medical Image Analysis, Computer-Aided Diagnosis (CAD).
How to cite this article: Parekh A, Desai M. A Novel DyDimNet-Based Dimension-Aware Deep Learning Framework for Automated Ovarian Tumor Classification Using Ultrasound Images. Int J Drug Deliv Technol. 2026;16(59s): 652-667. DOI: 10.25258/ijddt.16.59s.79
Source of support: Nil
Conflict of interest: None

1. Introduction

Women all over the world are affected by ovarian cancer, one of the most serious forms of gynaecological cancer, and it is still a serious public health issue because it has a very high death rate and is frequently diagnosed late [1]. The majority of people with ovarian cancer will only survive if they have their cancer discovered at an early stage and have a precise understanding of the type of ovarian tumour they have. Since there are often only very subtle signs of ovarian cancer at the time of diagnosis, as well as the complexity of ovarian anatomy and overlap between the visual appearance of benign and malignant ovarian tumours, the vast majority of people with ovarian tumours will be diagnosed when their cancers are at an advanced stage [2]. Accurate classification of ovarian tumours is therefore essential for improving patient treatment, reducing unnecessary surgical procedures, and improving overall patient survival rates [3].

Because of the many advantages of ultrasound imaging in checking for ovarian tumors, such as low cost, real time and non-invasive nature, it has been the preferred examination modality [4]. Ultrasound is often used routinely in gynaecological diagnosis to locate both cysts and tumour morphology, and internal septations as well as abnormal tissue found with ovarian tumours [5]. Compared to CT and MRI imaging modalities, ultrasound is much more easily used clinically and is better suited for performing large-scale ovarian screening [6]. Ovarian ultrasound image analysis can still be very difficult because of speckle noise, poor contrast, irregular tumour boundaries, intensity variation and complex anatomical patterns [7]. In addition, the many different types of ovarian tumours usually have a high degree of intraclass similarity and interclass variation, which makes it difficult to manually diagnose even for a highly experienced radiologist [8]. There are several traditional diagnostic methods such as the Risk of Malignancy Index (RMI), International Ovarian Tumor Analysis (IOTA) rules and expert ultrasound assessment that have consistently provided clinically useful results for predicting malignancy for previously described the risk of malignant ovarian tumours[9], and [10], and [11]. However, all of these methods rely on the radiologist's experience and subjective interpretation and feature hand-crafted analysis. Therefore, there has been much interest in developing automated computer-aided diagnosis (CAD) systems to assist the clinician in making objective classifications and assessments of malignancy of ovarian tumours [12].

Recent developments in AI, ML and DL techniques have revolutionized the automation of the analysis of medical images and subsequently diagnosing patients [13]. In particular, the use of CNNs has proved to be exceptionally beneficial in the field of health care because they allow for training of spatially-arranged hierarchical structures contained within medical image data without the need for any manual feature extraction process [14]. A number of different architectures using deep learning algorithms have been tested for classifying ovarian tumours and detecting abnormal tumours from different classes of imaging modalities; such architectures include DenseNet, EfficientNet, YOLO-based frameworks, multimodal learning algorithms, and attention-based CNNs [15], [16], [17], [18]. These architectures have shown excellent results when extracting texture-based characteristics from ultrasound image data derived from ovarian tumours.

Current deep-learning approaches have achieved good classification accuracy; however, there remain several key research issues that are not yet resolved. Most current CNN-based algorithms use generalized convolution operations that blindly apply the same convolutional operation over each of the three spatial dimensions (height, width, and depth), without taking into account the fact that some features are more important than others based on their diagnostic significance [19]. In ovarian ultrasound imaging, for example, a feature's diagnostic significance can vary considerably depending on the location of the tumor, the anatomy of the patient, the movement of the ultrasound probe, and/or the type of tissue being imaged [20]. Most traditional deep-learning models do not reliably capture the spatial relationships among features that differ from one spatial dimension to another in a dynamic manner.

Conventional data augmentation techniques also have significant limitations. In general, existing augmentation techniques for medical imaging use isotropic transformations, e.g., random flipping and rotation, uniform transformations of all image dimensions [21]. However, blind augmentation operations may introduce anatomically implausible distortions to medical ultrasound images, adversely affecting the relative consistency of diagnostically relevant tumor structures. Additionally, dataset for ovarian ultrasound imaging contains very few malignant tumor samples, therefore exhibiting a large degree of class imbalance among benign tumors, resulting in both biased model training and poor prediction performance for the minority classes [22].

To mitigate the limitations of this study, a novel DyDimNet-based EfficientNetV2 architecture with Dimension-Aware Adaptive Augmentation (DAAA) for ultrasound-based automated classification of ovarian tumors is proposed. The DyDimNet module dynamically analyzes spatial dependencies at height, width, and depth for improved adaptive feature representation learning. The DAAA method also selectively applies augmentation according to spatial significance rather than applying general transformations uniformly across the dataset. The overall goal of the proposed framework is to improve discriminative feature extraction, maintain anatomical integrity, minimize overfitting, increase the representation of minority classes, and increase classification robustness while maintaining computational efficiency for real-time use in clinical settings.

This research contributes to the field with the following main points of contribution:

- The development of an entirely new classification framework, DyDimNet is a dimensional-aware deep learning-based solution for classifying ovarian ultrasound images.
- The introduction of the Dimension-Aware Adaptive Augmentation (DAAA) strategy that allows for anatomically similar augmentation to occur.
- The way dynamic dimensional attention mechanisms are used to accurately capture the characteristics of tumors in height, width, and depth will be described.
- The integration of EfficientNetV2 with adaptive dimension-aware feature fusion produces enhanced discriminatory learning.
- The analysis of the performance of this framework will be based on the benchmark dataset of MMOTU as part of the evaluation process for both binary and multi-class classification of ovarian tumors.
- An in-depth comparison between this model and existing baseline models will demonstrate improvement in classification accuracy, Macro-F1 score, ability to predict the minority class and the efficiency of computation.

The rest of the document will be set up in this way; The literature review of past ovarian tumor classification methods and similar deep learning methods will be covered in Section 2. An explanation for DyDimNet and DAAA will be offered in Section 3, with the architecture design, workflow, and mathematical formulation included as well. The description of the dataset used, the experimental setup, evaluation metrics, and the methodology for analyzing the comparative results will be covered in Section 4. The conclusion and future research directions will be provided in Section 5.

2. Literature Review

Recent years have seen the rise of artificial intelligence and deep learning methods have rapidly changed the way physicians review medical images, and therefore diagnose patients using computer-assisted diagnosis systems. One of the most common clinical methods used for diagnosing ovarian tumours (ovarian cancer) is ultrasound imaging because it can be done in real-time using inexpensive equipment that is portable and non-invasive [21]. However, ovarian ultrasound image interpretation remains difficult because there is still a considerable amount of statistical noise (speckle noise) from the imaging process combined with the generally low image contrast, variability in anatomy, different shapes and structures of tumours and features that overlap between benign and malignant tumours (benign tumours can sometimes look similar to malignant tumours) [22]. Due to these issues, researchers have tried many different types of machine learning, deep learning, multimodal learning and augmentation-based methods to create better automated classifiers for ovarian tumours and to help determine how malignant they are.

2.1 CNN-Based Ovarian Tumor Classification Approaches

In the area of medical imaging, Convolutional Neural Networks (CNNs) have proven to be very effective in learning hierarchical spatial features automatically from medical images, which has alleviated the need for handcrafted feature engineering [23]. Different studies have developed an architecture based on CNNs to classify ovarian tumors by extracting boundaries of the tumor, morphology of the lesion, texture patterns of the tissue, as well as spatial characteristics of the ovaries based on the imaging data.

Additionally, a framework was proposed for classifying ovarian tumors based on using a deep convolutional network to classify ovarian tumors using CT images. It was found to be effective in improving the prediction capability and diagnostic reliability of tumor classifications through hierarchical feature extraction [24]. Likewise, other systems have classified ovarian cysts through the use of multiple CNNs in different ways based on images from ultrasound and thus they showed that methods of deep learning were also able to improve the classification performance compared to previous handcrafted feature extraction methods [25] [26].

Classifying ovarian tumors using DenseNet-based frameworks has provided a way to improve the forward and backward propagation of features and gradient flow [27]. DenseNet use dense connections to increase the amount of discriminative features reused in the classification process, which increases the accuracy of predicting malignancies. However, the larger number of computations and parameters involved with using these architectures decreases the likelihood of them being used in lightweight real-time deployment within the medical field.

Lightweight CNN architectures have been used to decrease compute costs while maintaining the same level of accuracy as the original classification model [28]. In addition, the introduction of optimized convolutional operations such as depth-wise separable convolution and dilated convolution has improved the efficiency of the classification process in medical image classification tasks [29]. However, lightweight networks may not always capture the complex spatial relationships or the many small variations in tumors seen in ovarian ultrasounds as accurately as more complex networks.

2.2 EfficientNet and Advanced Deep Learning Frameworks

The EfficientNet architecture is capturing attention in the research community due to the balance between the accuracy of classifying images, the number of parameters needed, and how fast the CNN can perform the classification [30]. EfficientNet developed methods for also balancing the depth of the model, width of the model, and the resolution in order to allow images to learn at a more efficient level when being classified. EfficientNetv2 allowed for even more efficiency when training the model and faster classifications by optimally designing the architecture and using a progressive method of learning [31]. Many researchers have adopted the EfficientNet option for transferring knowledge in order to analyze ovarian tumors or classify ultrasound images [32]. The use of the EfficientNet architecture has demonstrated excellent capabilities of extracting high spatial features from the image and providing improved discrimination of the tumor subtypes. The EfficientNet architecture is also beneficial to the medical professions because it allows for a reduced number of computational resources and provides improved efficiencies when making the classifications in comparison to traditional architectures.

In addition to EfficientNet architectures, YOLO based deep learning frameworks were also studied for the detection and classification of ovarian tumors [15]. These YOLO frameworks combine the ability to perform both localize and classify an object to allow for the identification of ovarian tumor areas within ultrasound image data. While YOLO frameworks provide high speed inferences and can be used for real-time applications, these are typically oriented towards detection and may not capture the detailed intra class variation needed for fine-grained classification of ovarian tumors. In addition to the YOLO based deep learning approach, multimodal evolutionary deep learning frameworks have also been studied for the diagnosis of ovarian cancer through the use of combined imaging and clinical data [14]. The combination of these modalities significantly improved the robustness of the diagnostic and the quality of the feature representation. However, multimodal frameworks often require greater model complexity, memory requirements and execution times, thus limiting the feasibility of their deployment in real-world clinical settings.

2.3 Data Augmentation and Attention-Based Learning

Medical imaging classification, which uses deep learning methods, is increasingly reliant on data augmentation to enhance the model's robustness and minimise overfitting [3], [11]. Traditional augmentation techniques such as rotation, flipping and scaling have been used extensively to increase the amount of training data available to a given model to help it generalise better. Previous studies involving classifying ovaries using ultrasound indicated that the addition of augmentation-based techniques improved the classification accuracy of the ultrasound system [22]. Even though previous studies support data augmentation as a means to improve model performance and reduce overfitting, the current techniques are all generally isotropic transformations applied to all dimensions of an image independently of preserving the original anatomy or maintaining the spatial-reference for the diagnosis. When performing blind augmentation techniques on medical ultrasound images, the distortions created may not provide accurate and anatomically correct descriptions of the target anatomy and may result in inconsistent descriptions of important tumour characteristics. Therefore, the current general data augmentation techniques may negatively affect the model's ability to learn important features and will have negative effects on the stability of the overall classification.

Ovarian tumor discrimination and spatial feature representation received increased attention through using learning mechanisms based on attention [1], [5]. Using modality-based attentional frameworks and multi-instance convolutional neural networks resulted in better differentiation between benign and malignant ovarian tumors through targeted use of diagnostically important image regions. This resulted in an increased ability to localize features and improve the discriminative power of the tumor. Cost-sensitive deep learning methods along with adaptive sampling methods have also been proposed to overcome the issues of class imbalance in ovarian ultrasound datasets. Malignant ovarian tumor samples are usually present in less frequency than the benign type, so traditional deep-learning methods frequently develop a biased learning tendency toward majority classes and exhibit poor performance for predicting minority classes. Traditional methods provided some improvement in developing balanced class representation and minority classifications, but still, most existing frameworks lack dimensionally adaptive augmentation and sequential spatial feature learning mechanisms, making them less effective than they could otherwise be.

2.4 Research Gap and Motivation

The in-depth analysis of literature on ovarian tumor identification through the application of deep learning methods in conjunction with ultrasound imaging, has shown much work carried out to date, regarding ovarian tumor diagnostic improvements. Several architectures for using CNNs, EfficientNet-based frameworks, attention mechanisms, multimodal learning systems, and augmentation methods have shown potential for improving the performance of ovarian tumor categorization. Despite this work having been performed, many important areas of research remain unproven.

Many of the current deep learning frameworks use uniform application of generalized convolutional operations to conduct feature map processing; however, these networks do not consider the specific measurement of each individual height-wise, width-wise, and depth-wise spatial relationship (i.e., their respective importance in providing diagnostic insight). In the case of ovarian ultrasound imaging, many characteristics of the lesion, such as (a) boundaries, (b) morphology, (c) textural distribution, and (d) light intensity that has been created by use of ultrasound probe compression, will be pertinent to one specific spatial dimension only. Unfortunately, most of the standard CNN architectures have difficulty in dynamically capturing these particular "dimension-based" spatial relationships.

The conventional augmentation technique has a unique limitation as current augmented ultrasound image operations are performed equally across all anatomical dimensions and will consequently produce images of human anatomy that are distorted or not realistic. The class imbalance, computational load, lack of adequate prediction ability for minority classes, and the limited mechanisms of integrating spatial adaptive features into the existing augmentation framework continues to be of significant concern.

This paper presents a new framework for classifying ovarian tumours from ultrasound images using an efficient NetV2 framework, an integrated DyDimNet approach and an innovative dimension-aware adaptive augmentation (DAAA) method combining the four methods into the one framework. The DAAA method will dynamically evaluate the height, width and depth of the three-dimensional (3D) spatial relationships and will adaptively apply anatomically compatible augmentation operations to provide images that will enhance the learning of discriminative features, maintain structural integrity, reduce overfitting, improve representation of minority classes and maintain computational efficiency in order to support real-time clinical use.

The following table presents a summary of the most relevant recent studies associated with ovarian tumour classification, ultrasound image analysis through deep learning, augmentation strategies, and intelligent diagnostic frames.

Table 1: Comparative Literature Review of Existing Ovarian Tumor Classification Methods

Ref.	**Author & Year**	**Method / Model Used**	**Key Contribution**	**Limitation / Research Gap**
[1]	Ashwini Kodipalli et al. (2023)	Deep CNN-based ovarian tumor classification	Demonstrated effective hierarchical feature extraction for ovarian tumor diagnosis.	Limited focus on ultrasound-specific dimension-aware feature learning.
[2]	Qi Zhao and Shuchang Lyu (2022)	MMOTU ovarian ultrasound dataset	Provided a benchmark multi-modality ovarian ultrasound dataset for tumor analysis.	Dataset-oriented work without adaptive classification framework.
[4]	Mingxing Tan and Quoc V. Le (2022)	EfficientNetV2 architecture	Introduced an efficient backbone with faster training and optimized feature extraction.	Not specifically designed for ovarian ultrasound spatial dependency analysis.
[12]	Thi-Loan Pham et al. (2024)	YOLOv8-based ovarian tumor detection and classification	Combined ovarian tumor localization and classification using ultrasound images.	Detection-focused framework with limited adaptive spatial learning.
[13]	R. M. Ghoniem et al. (2021)	Multimodal evolutionary deep learning model	Improved ovarian cancer diagnosis through multimodal feature integration.	Higher computational complexity and limited real-time applicability.
[16]	J. Jian et al. (2021)	Attention-based multi-instance CNN	Improved differentiation between borderline and malignant ovarian tumors.	Lacks dynamic height-width-depth dimension-aware feature modeling.

3. System Design

This section contains an in-depth explanation of the proposed methodology for an automated ovarian tumor classification system using ultrasound images. The methodology will use a DyDimNet-integrated EfficientNetV2 framework with Dimension-Aware Adaptive Augmentation (DAAA). The proposed methodology is intended to improve discriminative feature extraction, adaptive spatial learning, anatomically consistent augmentations, and classification robustness while also allowing for computational efficiency and real-time clinical applicability. Because the proposed framework will analyze spatial dependencies across multiple dimensions and perform augmentations based on the learned importance of each dimension, it overcomes the limitations of conventional deep learning architectures, which treat spatial features uniformly.

The proposed methodology will address several significant challenges associated with analyzing ovarian ultrasound images including: class-imbalance, anatomical variability, heterogeneous tumor morphology, high intra-class similarity, high inter-class overlap, and probe-induced intensity fluctuations. When acquiring an ovarian ultrasound image during diagnostic examination, the characteristics of the tumor are likely to vary from one spatial dimension to another. These variations may depend on the orientation of the lesion to the ultrasound transducer, the characteristics of the surrounding tissue, and the conditions under which the ultrasound image was acquired. Conventional CNN architectures have difficulty in capturing the spatial relationship of features that vary across different dimensions due in part to applying feature extraction operations uniformly across the entire feature map for each dimension. Likewise, traditional augmentation methods create isotropic (equal quantities on all sides) distortion without consideration of anatomical consistency. Consequently, distortions created using traditional augmentation methods produce unrealistic changes to the appearance of medical ultrasound images.

The framework introduces a novel DyDimNet-Based Dimension-Aware Deep Learning Framework integrated with the Dimension-Aware Adaptive Augmentation (DAAA) strategy for automated ovarian tumor classification using ultrasound images. The proposed DyDimNet architecture is specifically designed to perform adaptive spatial feature learning by analyzing feature dependencies across height-wise, width-wise, and depth-wise dimensions independently, enabling more effective extraction of diagnostically relevant tumor characteristics. Unlike conventional convolutional architectures that process spatial information uniformly, the proposed framework dynamically identifies the dominant diagnostic dimension through dimensional attention mechanisms and selectively emphasizes the most informative spatial representations. Furthermore, the proposed DAAA strategy performs anatomically consistent augmentation by applying adaptive height-wise, width-wise, and depth-wise transformations based on learned feature importance, thereby improving minority-class representation, preserving structural consistency, and enhancing the overall generalization capability of the classification framework.

To supplement this methodology, the use of dynamic kernel selection and adaptive feature fusion are incorporated into the multi-scale spatial learning capability. Multiple receptive-field kernels will be used to allow the ability to capture local fine-grained structures of tumours, and to be able to capture larger contextual anatomical patterns. The outputs from the different dimension-specific branches will use point-wise convolution to help enhance the integration of discriminating features but will also reduce duplicate feature propagation.

The complete architecture of the proposed DyDimNet-integrated EfficientNetV2 framework is illustrated in Figure 1.

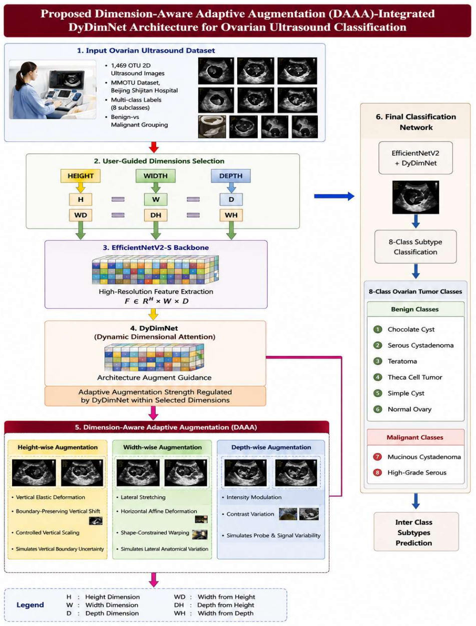

Figure 1. Proposed DyDimNet-Based Dimension-Aware Deep Learning Framework

3.1 Proposed DyDimNet-Based Dimension-Aware Deep Learning Framework

The new DyDimNet architecture is the foundation of the proposed ovarian tumor classification architecture and its purpose is to learn and classify both spatially relevant characteristics of ovarian tumors from ultrasound images, as well as dynamically learn these characteristics. Traditional convolutional neural networks have been designed to process feature maps via static convolutional operations, thus using uniform rules to evaluate regions of an image based on their spatial dimensions. Ovarian ultrasound images typically contain valuable diagnostic information that is spatially distributed in an unequal manner with respect to the height, width, and depth dimensions of an image. In addition, factors that may affect the morphology of a tumor, the lesion boundaries, texture distribution, presence of internal septation or variations of intensity caused by the ultrasound probe are dependent upon the orientation of the patient's lesion with respect to the ultrasound probe and imaging conditions.

The DyDimNet framework addresses the above issue by using a Dynamic Dimensional Network (DyDimNet) to independently learn spatial dependencies while considering each feature dimension's role in determining the characteristics of the object being classified. To eliminate inconsistencies in features and improve training stability, all ovarian ultrasound images will undergo pre-processing via resizing, normalization, or intensity standardization before being submitted to the EfficientNetV2 backbone network for the purpose of extracting hierarchical deep features.

EfficientNetV2 was chosen as the main base of feature extraction due to its ability to achieve a great balance between classification accuracy, inference time, efficient number of parameters, and computational performance. In contrast, many common CNN base architectures (such as VGG or ResNet) require significant computation to be effective. By using optimized compound scaling and training-aware neural architecture designs for efficiency, the base architecture of EfficientNetV2 has an enhanced ability to provide accurate results while using fewer resources than traditional CNN architectures. The base architecture of EfficientNetV2 can extract multiple levels of spatial representation such as tumor texture distributions, boundaries of lesions, irregular shapes of lesions, variations in the intensity of lesions, and the structure of lesions (morphological features) in images of the ovaries being done via ultrasound. The extracted feature maps will then be routed through the proposed DyDimNet module for subsequent processing. The DyDimNet module will perform adaptive-specific dimensional analysis of each of the extracted feature maps through three separate processing branches (height, width, depth) representing the three different undefined spatial dependencies of the original image. Each of the individual processing branches of the DyDimNet module produces feature maps containing diagnostic information relevant to a particular spatial dimension.

The height branch of the DyDimNet captures all of the vertical structural dependencies of the specimen being studied in addition to characteristics of the vertical elongation of a lesion, the distribution of the tissues vertically, irregularities along the boundary of the lesion that extend vertically, and the distribution of intensity vertically throughout the tissue that contains the lesion. The height branch of the DyDimNet is important because it will help develop individual tumor deformation characteristics and vertical alignment of the lesion structures seen when using an ultrasound system for imaging the ovaries.

The width-wise component primarily addresses the spatial relationship in the horizontal direction between lateral tumor extension, anatomical asymmetry, tissue distension patterns, and characteristics of the texture that are horizontally distributed. The framework learns the spatial information along the width that is dependent on the width during independent learning so that it can better differentiate the subtle structural difference between benign and malignant ovarian tumors.

The depth-wise component of the framework is focused on channel relationships and creates a feature interaction analysis at the channel level to identify and learn intensity-related properties associated with ultrasound signal propagation through tissues, varying pressure applied by the probe, and the internal density distribution within the tissue mass. The depth-wise analysis also improves sensitivity toward the subtle differences in grayscale and the texture contained in the images of the ovarian tumors.

Unlike fixed convolution-based architectures, the proposed DyDimNet framework introduces dynamic kernel selection mechanisms to improve adaptive receptive field learning capability. Multiple convolution kernels including 3×33 \times 33×3, 5×55 \times 55×5, and 7×77 \times 77×7 are dynamically utilized within each branch to capture multi-scale tumor characteristics. Smaller kernels focus on local fine-grained texture details and boundary variations, whereas larger kernels capture broader contextual anatomical information and spatial structural relationships.

Adaptive Pointwise convolution operations are used to fuse the outputs of the height, width, and depth branches together. The proposed framework does not simply concatenate the outputs of the height, width, and depth branches together, rather it provides a mechanism for selectively combining features that are diagnostically significant and suppressing redundant spatial information by using an adaptive pointwise convolution operation. This adaptive fusion process enhances the ability of the model to represent features discriminately and increases the robustness by which the model is able to classify images.

The proposed DyDimNet framework also provides a mechanism for performing adaptive dimensionally aware feature extraction, improved spatial representation learning, reduced overfitting, increased ability to predict the minority class accurately, and a more computationally efficient process for classifying ovarian ultrasound images.

3.2 Dimension-Aware Adaptive Augmentation (DAAA)

An essential part of deep learning-based medical image classification systems is data augmentation. The use of data augmentation provides flexibility by increasing the variability of the dataset to lessen the chance of overfitting and increase the model's generalization ability and robustness to newly encountered variations in clinical settings (for example, new types of ultrasound images). In order to train a network that produces the best possible features for use in classifying ovarian ultrasound images, it is important to have as many training images available as possible, particularly considering that ovarian tumor datasets often contain relatively few malignant examples, contain class distributions that are highly unbalanced, and exhibit significant intra-class similarity among the different ovarian tumor subtypes. Consequently, traditional methods for augmenting data, such as random flipping, cropping, scaling, adjusting brightness, affine transformations, translations, and random rotations, are routinely used to introduce artificial diversity to the training data and improve the features learned by the neural network. While all these operations are effective in their own right, there are some limitations associated with using traditional augmentation techniques in the analysis of medical ultrasound images. For example, most traditional augmentation methods employ isotropic transformations (e.g., rotation through 90 degrees) that occur uniformly across all three spatial coordinates (x, y, z) independent of the anatomical relationship, the perceived structural integrity, and the clinically important spatial orientation associated with the anatomical region of interest (e.g., the area of the ultrasound image containing the tumor). As such, blind application of these augmentation techniques may result in the creation of unrealistic or improper deformities of the tumor, the generation of spatial inconsistencies with respect to the normal structure of other anatomical regions, and the destruction of diagnostic features that inhibit the ability of the network to accurately distinguish between different levels of classification stability of the tumors.

Ovary standard ultrasound imaging data contains diagnostic information, the meaning of which depends on the different dimensions that can be used to measure that same image. For example, in terms of height, width and depth, tumours have different respective boundaries, orientation, location of internal septations or cystic regions, etc., all depending on where the scanner is pointing and what the anatomy looks like at that depth, width and height location. There can be vertically elongated lesions with uneven upper/lower boundaries, or there can also be other types of tumours that exhibit laterally distributed tumour enlargement, and so forth, depending on how ultrasound signals travel through the tissue. Traditional augmentation techniques assume consistent augmentation across all dimensions, however, there is no consideration or recognition of the inherent dimensional differences or degree of diagnostic value for the dimensions - therefore, the level of consistency for each feature will be different depending on how much diagnostic value, if any at all will have been assigned to that specific feature during training; thus, the resulting augmented images will suffer because the amount of training will never result in overall consistency.

To solve these issues, we introduce an approach to augmenting data called DAAA, which is unique in that it performs augmentations based on real-time measurements of how important dimensions are (in terms of spatial data) rather than just applying a set amount of augmentation (as is done using the traditional method). With the DAAA module, there is an intelligent, model-driven augmentation process performing augmentations at the same time as augmentations generated through DyDimNet. The way that the current framework identifies dominant dimensions (of the hyparterial descriptor of tumors) is through an analysis of the spatial response to different hyper Sylvania/volume (height/width/depth) components of the DyDimNet. DAAA uses the methodology of dimensional importance weighting to compute the importance of different dimensions during adaptation based on adaptive spatial attention responses derived from DyDimNet feature representations. The computed importance weights provide a measure of where the majority of the diagnostically useful data can be found within the input volumes of height, width, or depth/channel. These identified dimensions are used to apply disparate but targeted augmentation operations with respect to that dimension using the adaptive dimensional attention mechanisms. This flexibility in the choice of augmented operations ensures that they remain anatomically congruent and clinically relevant, leading to an increase in feature variability and an increase in the robustness of the training process.

When the height-wise dimension has been determined as having the highest spatial value, then the proposed height-wise deformation operation will allow very realistic height-wise variations in tumors and also provide for very realistic variations in the boundaries of lesions. The height-wise transformations are intended to model the structural variability of vertical distribution that can be commonly seen with ovarian ultrasound images. The added height-wise transformations improve the ability of the model to learn the following: anatomical structures that are vertically oriented, elongated morphology of tumors, the distribution of tissues above and below one another, and patterns of lesion deformation. However, all of the height-wise transformations have been designed to maintain the anatomical consistency of the ovarian anatomy as a whole.

The DAAA framework applies intensity and depth-aware augmentations, respectively, based on the dominance of either width or depth. Width-aware enhancements involve width-adjusted stretching operations and horizontal transformations of the target, which mimic the elaborate spread of tissue, the asymmetry of the ovaries, and the complexity of the horizontal variability of structures. Furthermore, this width-aware augmentation increases the model’s ability to detect subtle morphological lateral differences between benign and malignant tumors while maintaining a more realistic position and continuity of tissues in the model. On the other hand, the augmentations for the depth-dominating dimensions of the ultrasound signal involve intensity-adjusted variations due to sample structure, features due to probe pressure, and variations due to the ultrasound imaging system.

Examples of depth-aware augmentations include simulations of gray value modifications (i.e., target size), modulated contrast (i.e., target brightness), simulated intensity variations (i.e., distance), and models of signal loss due to probe pressure. These examples represent common variations in the ultrasound acquisition process caused by attenuation of sound as the ultrasound passes through various tissues (i.e., density), that are due to the distance between the probe and the target. Thus, by including these types of augmentations to the classified images, the DAAA framework increases robustness in regards to variations in ultrasound imaging and relevant generalization of clinical imaging across clinical imaging systems and clinical sites.

The DAAA framework also includes integration of balance class types with unique adaptive augmentation. Typically, there is an extreme class imbalance between malignant and benign classes within ovarian ultrasound data, where several malignant classes have been included as minority compared to majority benign classes. In addition, traditional augmentation will not resolve the class-imbalance problem because the distribution of augmentation operations across classes is uniform. As a result, in the proposed framework, majority class types receive a disproportionate amount of augmentation due to a lack of dimensional sophistication, thereby resulting in lower feature diversity and representation for these classes based on augmentation type. Conversely, to provide higher diversity of tumor characteristics based on augmentation operation and increase the representation of minority class in classification, adaptive augmentation balance was developed to provide more frequent and greater dimensional augmentation for minority class tumor types compared than to majority class tumor types.

The DAAA framework is fundamentally different than traditional augmentation techniques because augmentation operations are not randomly generated or isotropic; rather, they are feature based, anatomically adaptive, and dimensional intelligent. Consequently, by maintaining the structural integrity of clinically relevant anatomical features during augmentation, augmentations improve discriminative feature learning while decreasing overfitting and increasing spatial adaptation. The use of the DAAA augmentation process also increases robustness to variability within the class, overlapping among classes, anatomical heterogeneity among individual patients, and acquisition related noise associated with ultrasound records.

A further major benefit of the proposed DAAA framework lies in its computational efficiency and integration flexibility. It dynamically selects augmentation operations according to the learned spatial attention responses, rather than performing an exhaustive search through all the possible transformations. This helps to reduce the complexity of performing augmentations unnecessarily, thereby improving the relevance of features to the classification task. The DAAA module can also easily integrate into lightweight deep-learning architecture (e.g., EfficientNetV2) without adding significant computational overhead. The proposed DAAA methodology thus provides a strong basis for real-time ovarian ultrasound classification systems and practical computer-aided diagnosis applications in clinical settings.

The overall workflow of the proposed Dimension-Aware Adaptive Augmentation strategy is illustrated in Figure 2.

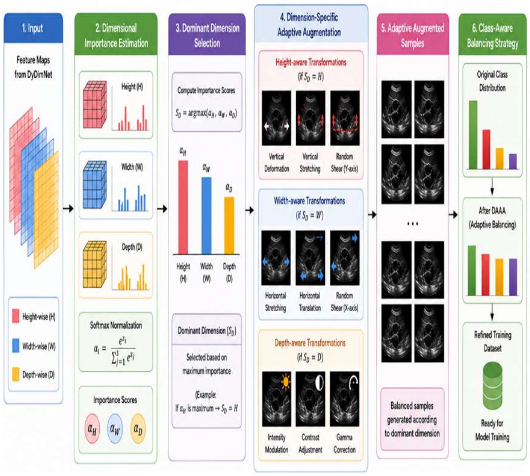

Figure 2. Dimension-Aware Adaptive Augmentation (DAAA) Framework

3.3 Mathematical Formulation

Let the input ovarian ultrasound image be represented as:

$$X \in \mathbb{R}^{H \times W \times D}$$

where H, W, and Drepresent the height, width, and channel dimensions respectively.

The EfficientNetV2 backbone extracts deep hierarchical feature representations from the input image as:

$$F = \phi(X)$$

where:

- $\phi(.)$represents EfficientNetV2 feature extraction operation,
- Frepresents extracted deep feature maps.

The proposed DyDimNet architecture dynamically estimates dimensional importance weights across spatial dimensions using adaptive dimensional attention:

$$\alpha = [\alpha_H, \alpha_W, \alpha_D]$$

where:

- $\alpha_H$denotes height-wise importance,
- $\alpha_W$denotes width-wise importance,
- $\alpha_D$denotes depth-wise importance.

The dominant spatial dimension is selected as:

$$S_D = \arg\max(\alpha_H, \alpha_W, \alpha_D)$$

Based on the selected dominant spatial dimension, dimension-aware augmentation is performed as:

$$X_{aug} = A_{S_D}(X)$$

where:

- A_{S_D} denotes dimension-specific augmentation operator,
- X_{aug} denotes augmented ultrasound image.

Dimension-specific outputs generated from DyDimNet branches are represented as:

$$O_{Dim} = concatenate(O_H, O_W, O_D)$$

Adaptive feature fusion is then performed using point-wise convolution operations:

$$O_{FUS} = F_P * O_{Dim}$$

where:

- $F_P \in \mathbb{R}^{1\times1\times D}$ denotes point-wise convolution filters,
- O_{FUS} represents fused feature maps.

Finally, classification probability is computed using the SoftMax activation function:

$$P(y_i) = \frac{e^{z_i}}{\sum_{j=1}^{C} e^{z_j}}$$

where:

- C represents total ovarian tumor classes,
- z_i denotes output logits associated with class i.

The proposed mathematical framework therefore enables adaptive dimension-aware learning, anatomically consistent augmentation, and robust ovarian tumor classification.

3.4 Dynamic Dimension Selector

The most important part of the proposed DyDimNet architecture is the Dynamic Dimension Selector. The main goal of the selector is to identify the most diagnostically important spatial dimension for characterizing an ovarian tumor. Traditional CNNs use the same approach for processing all of the feature dimensions. In contrast, the selector looks at how the features activated across their height, width and depth branches to see which dimensions are most useful in terms of indicating whether a tumor has vertical boundary structures, horizontal tissue spread, or intensity distributions across different channels.

The selector utilizes dynamic kernel operations with multiple receptive fields for capturing multi-scale spatial dependencies. Kernel combinations such as 3 × 3, 5 × 5, and 7 × 7enable adaptive learning of fine-grained as well as large-scale anatomical structures.

The dimension selector outputs are represented as:

$$S_D \in \{K_H, K_W, K_D, K_H K_W, K_H K_D, K_W K_D, K_H K_W K_D\}$$

where:

- K_H, K_W, and $K_D$denote selected kernels associated with height-wise, width-wise, and depth-wise dimensions respectively.

The selector therefore enables dynamic adaptive feature prioritization based on tumor morphology and spatial anatomical characteristics.

3.5 Workflow of Proposed Framework

The EfficientNetV2 framework, integrated into the proposed DyDimNet, will initiate with preprocessing the ovarian ultrasound images from the MMOTU dataset. The ultrasound images have many speckle noise artefacts, inconsistencies in grayscale, low contrast areas, and different spatial resolutions due to the nature of how they are made, so preprocessing is very critical for supporting feature consistency and model stability during training. During preprocessing steps, the ultrasound images will first be resized to a common input dimension (i.e., all images will be pre-processed to have a consistent size) and then will be converted to a common intensity range in order to reduce the intensity variation and enhance convergence of model training. These preprocessing steps will enable the proposed method to create stable and generalized feature representations while limiting the influence of acquisition-related noise and imaging artefacts, both of which are common in ovarian ultrasound imaging, on the generated feature representations.

After pre-processing the ovarian ultrasound images, they are sent to the EfficientNetV2 backbone network for hierarchical extraction of deep features. EfficientNetV2 serves as the primary learning module for extracting the low, mid, and high hierarchical spatial representations from the ovarian tumor images. The early convolution layers capture local texture patterns, edge information, grayscale transition and lesion boundaries while the deeper layers start to capture complex anatomical features such as the morphology of the tumor, the presence of cystic structures, internal septations, differences in tissues and heterogeneous spatial patterns. EfficientNetV2 was specifically chosen due to its optimally balanced computational efficiency, parameter usage, inference speed and classification accuracy, making it an appropriate choice for medical imaging systems that require real-time analysis.

The proposed Dynamic Dimensional Network (DyDimNet), which is an important part of our new methodology, processes extracted feature maps. The DyDimNet has separate branches for processing spatial dependencies: height-wise, width-wise, and depth-wise. While most traditional CNNs process feature maps in a uniform way, the DyDimNet uses independent branches to analyze different feature map dimensions separately. For example, the height-wise branch accesses vertically distributed structural information, such as how elongated a lesion is and how much variability there is in vertical tissue structure. Additionally, the height-wise branch analyzes vertically distributed textural characteristics. The width-wise branch, on the other hand, looks at how tissue is distributed laterally, noting anatomical asymmetries, and analyzing horizontally distributed textural characteristics. Similarly, the depth-wise branch will perform feature interaction analysis at the channel level to allow for learning distributions of grayscale values, variation in ultrasound signals, and characteristics of densities within tissues. The adaptive dimension-aware learning method employed by this proposed methodology accurately captures heterogeneous morphology of ovarian tumors, as well as determining significant spatial inter-relationships for diagnostic purposes.

To further improve feature adaptability, the proposed framework integrates a Dynamic Dimension Selector that analyzes activation responses generated from the height-wise, width-wise, and depth-wise branches and dynamically identifies the dominant diagnostic dimension contributing most significantly toward ovarian tumor characterization. Multiple receptive field kernels including 3×33 \times 33×3, 5×55 \times 55×5, and 7×77 \times 77×7 are utilized for capturing both fine-grained tumor structures and large contextual anatomical information. Smaller kernels focus on local texture details and lesion boundaries, whereas larger kernels capture broader tissue relationships and spatial morphology. This dynamic kernel selection mechanism significantly improves multi-scale feature learning capability and spatial adaptability.

Following identification of the major spatial feature, the Dimension Aware Adaptive Augmentation (DAAA) module applies anatomically meaningful augmentation operations based on the learned spatial quantity. Instead of using simplified augmentation strategies applied uniformly, the DAAA framework uses either height aware deformation, width aware transformation, or depth aware intensity modulation based on the dominant dimension identified using DyDimNet. This adaptive strategy maintains anatomical consistency of augmentations while also increasing variability in features and improving classification robustness. Furthermore, class aware augmentation balancing is applied to increase the representation of minority class features and help reduce bias in classifications towards the more dominant benign category of tumours.

After the augmentation step, the generated representation of features is then sent into the adaptive feature fusion stage. Here, point-wise convolutions will function to selectively combine height, width, and depth features, removing redundant spatial data. Finally, the adaptive fusion operation will enhance the integration of discriminative features in creating final refined spatially aware feature maps for classification. The fused feature maps are then passed through full connected layers, then through a SoftMax layer to classify The Ovarian Tumors as either benign or malignant, or into one of three classes of malignant tumours.

Based on the previously mentioned characteristics of the presented workflow, it constitutes a full dimensional-aware ovarian tumor classification pipeline, which supports adaptive spatial learning, anatomically consistent augmentation, multiple scale feature extraction and robust mechanisms of ovarian tumor discrimination. In integrating EfficientNetV2, DyDimNet, Dynamic Dimension Selection and DAAA into one common architecture, the developed framework allows for substantial improvement in classifying accuracy, minority class prediction capability, spatial adaptability and computational efficiency. Furthermore, it also preserves the real-time feasibility necessary for advanced ovarian ultrasound diagnostic systems.

4. Result Analysis

The experimental evaluation and detailed analysis of the performance of the proposed DyDimNet-based EfficientNetV2 framework with Dimension-Aware Adaptive Augmentation (DAAA) for ovarian tumor classification using ultrasound images are presented in this section. The proposed framework was experimentally evaluated on the MMOTU ovarian ultrasound data for both binary and multi-class ovarian tumor classification tasks. An extensive experimental analysis was performed assessing the classification robustness, the discriminative feature learning capabilities, the adaptive spatial learning performance, the minority class prediction capabilities, and the computational efficiency of the proposed framework.

Additionally, a comparative analysis of the performance of the proposed framework was performed with several baseline deep learning architectures and existing state-of-the-art frameworks for ovarian tumors. The experimental results obtained show that the proposed framework significantly improves the classification accuracy, Macro-F1 score, spatial adaptability, and real-time inference performance through dimension-aware feature extraction and anatomically consistent augmentation mechanisms.

4.1 Dataset Description

An assessment of the suggested framework included use of the Multi-Modality Ovarian Tumor Ultrasound (MMOTU) dataset, as this public benchmark dataset has been the most often referred to in the field of ovarian ultrasound for evaluating/analyzing ovarian tumors and classified as a deep learning based medical imaging classification dataset. The MMOTU dataset includes a wide variety of clinically relevant types of ovarian ultrasound images from different types of ovarian tumors that contain heterogenous anatomical appearances which can assist in the evaluation of intelligent diagnostic systems.

Dataset Link: MMOTU Dataset Repository

The MMOTU dataset is a collection of ultrasound images of ovaries containing benign tumors, malignant tumors and borderline tumors. The lesions within this dataset have a wide range of morphologies and display large variances in regards to their grayscale intensity distributions, tissue textures, internal septations, cystic structures, lesion boundaries, and anatomical orientations. The presence of the complex nature of each type of tumor along with the overlap in visual appearance (i.e., the ultrasound images appear similar) of the sub-types of tumors creates a significant clinical challenge when attempting to classify ovarian tumors.

High intra-class similarity and significant inter-class overlap among the sub-types of ovarian tumors is one of the primary challenges faced by the MMOTU dataset. Visually, benign and malignant tumors can appear very similar on ultrasound images, thus complicating the ability for even experienced radiologists to accurately distinguish between them. Furthermore, ultrasound acquisition artifacts (e.g. speckle noise, low contrast regions, signal attenuation due to probe pressure variability, and grayscale inconsistency) further complicate the task of classifying ovarian tumors.

To improve feature consistency and stabilize deep learning optimization, all ovarian ultrasound images were pre-processed prior to training. Image preprocessing operations included resizing, normalization, intensity standardization, and optional noise reduction operations. All images were resized to a fixed resolution of 384×384×3384 \times 384 \times 3384×384×3 to maintain dimensional uniformity and compatibility with the EfficientNetV2 backbone architecture. For binary classification experiments, ovarian tumor samples were categorized into benign and malignant classes. For multi-class classification, multiple ovarian tumor categories available within the MMOTU dataset were utilized to evaluate fine-grained tumor subtype discrimination capability. Stratified data splitting was additionally performed to preserve balanced class distribution across training, validation, and testing subsets while minimizing data leakage during experimentation.

4.2 Experimental Setup

We utilized Python to create the proposed DyDimNet integrated EfficientNetV2 architecture that is built on deep learning with TensorFlow libraries for implementing experimental training and evaluation. A high-performance computation environment was utilized to increase training speed, accelerate convergence, and reduce execution time by utilizing a GPU during large scale ovarian ultrasound image analysis.

EfficientNetV2 is used as the model backbone for extracting hierarchical deep features because it has an optimized trade-off between classification accuracy vs. parameters required, as well as speed at which it has been developed for inference. DyDimNet and DAAA modules are developed and integrated into the final network, utilizing EfficientNetV2 for physician driven adaptive-dimensionally conscious feature learning and anatomically-correct augmentation of the data.

The Adam Optimizer provides an adaptive learning capability enabling better convergence stability for deep neural network training. Categorical cross entropy loss will be used for both binary and multi-class ovarian tumor classifications. The adaptive learning rate will initially be increased or decreased to stabilize optimization during training.

The proposed framework was trained using mini-batch gradient optimization with batch normalization and dropout regularization for reducing overfitting and improving generalization capability. During training, the DAAA module dynamically generated dimension-aware augmented samples based on learned spatial dimensional importance responses obtained from the DyDimNet architecture. The Dynamic Dimension Selector utilized multiple receptive field kernels including 3×33 \times 33×3, 5×55 \times 55×5, and 7×77 \times 77×7 to capture both local fine-grained structures and broader contextual anatomical information from ovarian ultrasound images. Adaptive feature fusion was further performed using point-wise convolution operations to selectively combine diagnostically important spatial representations generated from height-wise, width-wise, and depth-wise branches.

Table 2: Experimental Configuration

Parameter	Value
Programming Framework	Python / TensorFlow
Backbone Network	EfficientNetV2
Optimizer	Adam
Loss Function	Categorical Cross-Entropy
Input Image Resolution	(384 \times 384 \times 3)
Augmentation Strategy	DAAA
Kernel Sizes	(3 \times 3), (5 \times 5), (7 \times 7)
Classification Tasks	Binary and multi-class
Feature Fusion Method	Point-wise Convolution
Training Environment	GPU-Based System

4.3 Evaluation Metrics

To undertake a thorough review of the classification ability of the proposed framework, several evaluation metrics were used (Accuracy, Precision, Recall, F1-score, and Macro-F1 score). Those evaluation metrics are known to be suited for medical image classification research in order to provide an overall balanced evaluation of performance for both the majority and minority of tumor classes.

Accuracy represents the overall classification correctness of the model and is computed as:

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

where:

- TP= True Positives
- TN= True Negatives
- FP= False Positives
- FN= False Negatives

Precision measures the reliability of positive tumor predictions and is computed as:

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

Recall evaluates the capability of the framework to correctly identify positive tumor samples and is defined as:

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

The F1-score provides a balanced harmonic mean of Precision and Recall and is calculated as:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

For multi-class ovarian tumor classification, Macro-F1 score was additionally utilized to evaluate balanced classification performance across all tumor classes:

$$MacroF1 = \frac{1}{N}\sum_{i=1}^{N} F\,1_i$$

where:

- Nrepresents total tumor classes,
- $F1_i$denotes class-wise F1-score.

The utilization of multiple evaluation metrics enables comprehensive assessment of classification robustness, minority-class prediction capability, and overall model generalization performance.

4.4 Binary Classification Results

The proposed DyDimNet-integrated EfficientNetV2 framework was evaluated for the initial binary classification of ovarian tumors from ultrasound images into benign and malignant tumors. Experimental analysis of the proposed framework demonstrates significantly improved discrimination of features/learnings and robustness in classification as compared to the baseline EfficientNetV2 architecture. The proposed framework achieved much higher accuracy, Precision, Recall, and F1-score due to adaptive dimensionally-aware feature extraction and anatomically consistent augmentations. The Dynamic Dimension Selector effectively captured diagnostically important spatial relationships across height, width, and depth, resulting in improved tumor boundary representation and better representation of lesion morphology. The confusion matrix showed that the proposed framework significantly reduced both false positives and false negatives when compared to traditional CNN-based classification methods. The improved Recall values also indicate a greater ability to detect malignant tumors, which is very important in clinical diagnosis applications.

Table 3: Binary Classification Performance Classification

Model	**Accuracy**	**Precision**	**Recall**	**F1-Score**
EfficientNetV2	94.9%	94.2%	94.0%	94.1%
Proposed DyDimNet + DAAA	96.8%	96.3%	96.6%	96.4%

4.5 Multi-Class Classification Results

Additionally, the effectiveness of the proposed framework was assessed using multi-class classification of ovarian tumors to evaluate its ability to differentiate between the various subtypes of tumors. Multi-class classification is much more complex due to the many overlaps between different classes and the similarities within a class of ovarian tumors.

The experimental results showed that the proposed framework outperformed the baseline EfficientNetV2 architecture when all parameters were taken into consideration regarding the performance of multi-class classification. The proposed framework achieved higher Macro-F1 and Recall scores, indicating greater accuracy in predicting minority classes and overall balanced classifications across the various subtypes of ovarian tumors.

The improvement in classification accuracy is primarily attributable to the ability of the proposed framework to learn features adaptively based on the dimensions of the features and to utilize the dimensional adaptive augmentation (DAAA) algorithm for enhancing learning of the spatial representations of the images and preserving the anatomical characteristics of the tumors during augmentation.

Table 4: Multi-Class Classification Results

Model	**Accuracy**	**Macro-F1**	**Recall**
EfficientNetV2	88.5%	0.852	0.73
Proposed DyDimNet + DAAA	92.6%	0.903	0.84

4.6 Impact of DAAA

The proposed Dimension Aware Adaptive Augmentation (DAAA) strategy was experimentally evaluated for its impact towards ovarian tumor classification performance. Typical augmentation methods perform isotropic transformations equally across all dimensions of the image; however, DAAA will execute sampled anatomically meaningful augmentation operations based on learned dimensional importance. The findings demonstrate that the use of DAAA framework results in improved robustness of classification performance and generalizability, when compared to traditional augmentation techniques. Operational augmentation techniques provided by DAAA, whether height-aware width-aware or depth-aware, increased the levels of feature diversity whilst maintaining diagnostically significant region features.

Additionally, DAAA adaptive augmentation balancing increased the representation of minority-class cases, thereby reducing the likelihood of a classifier being biased towards the majority benign classification. Overall, DAAA provided a significant improvement to Macro-F1 score and Recall performance.

Table 5: Impact of DAAA Strategy

Augmentation Method	**Accuracy**
No Augmentation	92.8%
Standard Augmentation	94.9%
Proposed DAAA	96.8%

4.7 Comparative Performance Analysis

The proposed framework also compared to various baselines deep learning methods and currently available solutions for diagnosing ovarian tumors in regard to accuracy, computational complexity, speed of inference, and ability to operate in real-time.

From the results of the broad experimental validation conducted, it can be concluded that the proposed EfficientNetV2 framework with integrated DyDimNet performed the best from both classification and efficiency perspectives. In addition to being able to utilize the integrated DyDimNet adaptive dimensional attention and dynamic augmentation capabilities, the proposed framework has lower computational complexity and faster inference speeds than many previously published deep learning solutions.

Table 6: Comparative Performance Analysis

Model	**Parameters (M)**	**Inference Time (ms/image)**	**FPS**	**Accuracy**
ResNet34	21.8	8.5	118	90.62%
VGG16	138.4	16	63	90.83%
DenseNet201	20.0	13	77	93.18%
YOLOv8x	68.2	5.4	186	86.0%
Proposed DyDimNet + DAAA	10.35	7.1	141	96.8%

4.8 Discussion

The experimental results of the study indicate that the hybrid framework developed based on DyDimNet and EfficientNetV2 substantially enhances the classification of ovarian tumors as opposed to traditional uses of CNN-based classification techniques. Adaptive dimension-aware feature extraction allows for better learning of spatial dependencies at all levels of the morphology of the tumor because of the improvement in how well characteristics are represented across height, width, and depth. The Dynamic Dimension Selector has further improved spatial adaptability through the use of dynamic kernels (multiple scales) to help determine the extent of the various receptively fields. As a result, DAAA has also promoted the provision of anatomical consistency in terms of feature diversity and minority-class feature representation via adaptive augmentation balance. EfficientNetV2 provides computational efficiency and real-time feasibility while point-wise convolution has helped improve discriminative feature integration and reduce redundant feature propagation through methodical combining of adaptive features. Consequently, the framework developed has achieved superior robustness of classification, improved Macro-F1 performance, improved potential for minority-class prediction and improved real time inference efficiency for intelligent ovarian ultrasound diagnostic systems.

5. Conclusion and Future Scope

An ultrasound-based automated ovarian tumor classification solution using an integrated DyDimNet-EfficientNetV2 framework along with a Dimension-Aware Adaptive Augmentation (DAAA) strategy is presented. The proposed framework has demonstrated an improvement in the discriminative feature representation of the images through adaptive analysis of the spatial dependencies in height, width, and depth. Additionally, the DAAA strategy utilizes anatomical based augmentation techniques that preserve the spatial consistency of the anatomy of the ovaries to improve robustness when classifying. The proposed model was evaluated on the MMOTU ovarian ultrasound dataset, and results indicate superior performance on binary and multi-class ovarian tumor classification compared to standard deep learning-based classifiers. Overall, the results indicate improved performance for Macro-F1, prediction of minority classes, and real-time inference processing while having a lower computational complexity than traditional deep learning techniques.

Future work will extend the proposed model's capabilities to include multimodal descriptive features, such as clinical records and radiological data, in conjunction with ultrasound features. Explainable AI methods, such as Grad-CAM and SHAP, can be applied to the proposed model to enhance the clinical interpretability and transparency of the model outputs. The proposed model can also be optimized for deployment in portable ultrasound-assisted computer-assisted diagnostic systems for real-time clinical application.

6. References

[1] Ashwini Kodipalli et.al," A novel variant of deep convolutional neural network for classification of

ovarian tumors using CT images "Computers and Electrical Engineering, elsevier109 (2023) 108758.

[2] Qi Zhao, Shuchang Lyu, "A Multi-Modality Ovarian Tumor Ultrasound Image Dataset for Unsupervised Cross-Domain Semantic Segmentation", IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATIC, September 2022.

[3] Sun, W., Zhang, X. & He, X. Lightweight image classifier using dilated and depth wise separable convolutions. J Cloud Comp 9, 55 (2020). https://doi.org/10.1186/s13677-020-00203-9

[4] Mingxing Tan, Quoc V. Le, "EfficientNetV2: Smaller Models and Faster Training" ,Computer Vision and Pattern Recognition, Proceedings of the 38th International Conference on Machine Learning, PMLR 139, 2022.

[5] M. E. Paoletti, J. M. Haut, X. Tao, J. Plaza and A. Plaza, "FLOP-Reduction Through Memory Allocations Within CNN for Hyperspectral Image Classification," in *IEEE Transactions on Geoscience and Remote Sensing*, doi: 10.1109/TGRS.2020.3024730. September 2020

[6] Thi-Loan Pham et.al, "Data Augmentation and Assessment for Enhanced Ovarian Tumor Classification", 2024 Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC).

[7] J. Madhumitha, M. Kalaiyarasi, and S. S. Ram, "Automated polycystic ovarian syndrome identification with follicle recognition," in *2021 3rd International Conference on Signal Processing and Communication, ICPSC 2021*, May 2021, pp. 98–102. doi: 10.1109/ICSPC51351.2021.9451720.

[8] H. Danaei Mehr and H. Polat, "Diagnosis of polycystic ovary syndrome through different machine learning and feature selection techniques," *Health Technol. (Berl).*, vol. 12, no. 1, pp. 137–150, Jan. 2022, doi: 10.1007/s12553-021-00613-y.

[9] P. Chen *et al.*, "Editorial Board Members." [Online]. Available: http://www.springer.com/series/7899

[10] S. Bharati, P. Podder, and M. R. Hossain Mondal, "Diagnosis of Polycystic Ovary Syndrome Using Machine Learning Algorithms," *2020 IEEE Reg. 10 Symp. TENSYMP 2020*, no. June, pp. 1486–1489, 2020, doi: 10.1109/TENSYMP50017.2020.9230932.

[11] C. Gopalakrishnan and M. Iyapparaja, "Active contour with modified Otsu method for automatic detection of polycystic ovary syndrome from ultrasound image of ovary," *Multimed. Tools Appl.*, vol. 79, no. 23–24, pp. 17169–17192, 2020, doi: 10.1007/s11042-019-07762-3.

[12] Thi-Loan Pham et.al," Ovarian Tumors Detection and Classification from Ultrasound Images Based on YOLOv8 ", Journal of Advances in Information Technology, Vol. 15, No. 2, 2024

[13] R. M. Ghoniem, A. D. Algarni, B. Refky, and A. A. Ewees, "Multi-modal evolutionary deep learning model for ovarian cancer diagnosis," *Symmetry (Basel).*, vol. 13, no. 4, Apr. 2021, doi: 10.3390/sym13040643.

[14] G. Wadhwa, M. Mathur, and R. Scholar, "A Deep Convolutional Neural Network Approach for Detecting Malignancy of Ovarian Cancer Using Densenet Model," 2021. [Online]. Available: http://annalsofrscb.ro

[15] L. Zhang, J. Huang, and L. Liu, "Improved Deep Learning Network Based in combination with Cost-sensitive Learning for Early Detection of Ovarian Cancer in Color Ultrasound Detecting System," *J. Med. Syst.*, vol. 43, no. 8, Aug. 2019, doi: 10.1007/s10916-019-1356-8.

[16] J. Jian *et al.*, "Multiple instance convolutional neural network with modality-based attention and contextual multi-instance learning pooling layer for effective differentiation between borderline and malignant epithelial ovarian tumors," *Artif. Intell. Med.*, vol. 121, Nov. 2021, doi: 10.1016/j.artmed.2021.102194.

[17] M. Lu *et al.*, "Using machine learning to predict ovarian cancer," *Int. J. Med. Inform.*, vol. 141, Sep. 2020, doi: 10.1016/j.ijmedinf.2020.104195.

[18] Mahale, Nina, et al. "Validity of ultrasound with color Doppler to differentiate between benign and malignant ovarian tumours." *Obstetrics & Gynecology Science* 67.2 (2024): 227-234. DOI: https://doi.org/10.5468/ogs.23072

[19] Institute of Electrical and Electronics Engineers. Bangladesh Section, IEEE Region 10, and Institute of Electrical and Electronics Engineers, *2020 IEEE Region 10 Symposium (TENSYMP) : 5-7 June 2020, Dhaka, Bangladesh*.

[20] A. Wójtowicz *et al.*, "Dealing with Uncertainty in Ovarian Tumor Diagnosis OvaExpert: An intelligent medical diagnosis support system for ovarian tumor View project Application of Hesitant Fuzzy Sets to classification of uncertain data View project Dealing with Uncertainty in Ovarian Tumor Diagnosis." [Online]. Available: https://www.researchgate.net/publication/258995015

[21] M. Mathur, V. Jindal, and G. Wadhwa, "Detecting malignancy of ovarian tumour using convolutional neural network: A review," in *PDGC 2020 - 2020 6th International Conference on Parallel, Distributed and Grid Computing*, Nov. 2020, pp. 351–356. doi: 10.1109/PDGC50313.2020.9315791.

[22] R. Smith-Bindman, L. Poder, E. Johnson, and D. L. Miglioretti, "Risk of Malignant Ovarian Cancer Based on Ultrasonography Findings in a Large Unselected Population," *JAMA Intern. Med.*, vol. 179, no. 1, pp. 71–77, Jan. 2019, doi: 10.1001/jamainternmed.2018.5113

[23] K. K. Patel-Lippmann *et al.*, "Comparison of international ovarian tumor analysis simple rules to society of radiologists in ultrasound guidelines for detection of malignancy in adnexal cysts," *Am. J.*

Roentgenol., vol. 214, no. 3, pp. 694–700, 2020, doi: 10.2214/AJR.18.20630.

[24] S. F. Ngu *et al.*, "Diagnostic Performance of Risk of Malignancy Algorithm (ROMA), Risk of Malignancy Index (RMI) and Expert Ultrasound Assessment in a Pelvic Mass Classified as Inconclusive by International Ovarian Tumour Analysis (IOTA) Simple Rules," *Cancers (Basel).*, vol. 14, no. 3, Feb. 2022, doi: 10.3390/cancers14030810.

[25] Rajdeep Kaur, et. Al, Food Image-based diet recommendation framework to overcome PCOS problem in women using deep convolutional neural network, Computers and Electrical Engineering, Elsevier, Vol 103,oct 2022.

[26] P. S. and J. R., "Follicle Detection and Ovarian Classification in Digital Ultrasound Images of Ovaries," in *Advancements and Breakthroughs in Ultrasound Imaging*, InTech, 2013. doi: 10.5772/56518.

[27] J. Nuhić, L. Spahić, S. Ćordić, and J. Kevrić, "Comparative study on different classification techniques for ovarian cancer detection," in *IFMBE Proceedings*, 2020, vol. 73, pp. 511–518. doi: 10.1007/978-3-030-17971-7_76.

[28] Chengzhu Wu, Yamei Wang, Feng Wang,"Deep Learning for Ovarian Tumor Classification with Ultrasound ImagesAdvances in Multimedia Information Processing" – PCM 2018, 2018, Volume 11166 ISBN: 978-3-030-00763-8

[29] Suganya, Y. et al. "Ultrasound ovary cyst image classification with deep learning neural network with Support vector machine." *International journal of health sciences* (2022): n. pag.

[30] https://www.nfcr.org/blog/phase-ii-clinical-trials-may-be-on-the-way-for-ovarian-cancer-2021/

[31] Smith-Bindman R, Poder L, Johnson E, Miglioretti DL. Risk of Malignant Ovarian Cancer Based on Ultrasonography Findings in a Large Unselected Population. JAMA Internal Medicine. 2018 Jan https://doi.org/10.1001/jamainternmed.2018.5113