



Research Article

A novel dynamic dimensional convolution network architecture for classification of ovarian tumor

Aarti PAREKH^{1,*}, Madhavi DESAI²

¹Gujarat Technological University, Gujarat, 382424, India

²Department of CSE, R.N.G. Patel Institute of Engineering, Gujarat, 394601, India

ARTICLE INFO

Article history

Received: 10 July 2025

Revised: 03 October 2025

Accepted: 12 March 2026

Keywords:

Convolutional Neural Network (CNN); Deep Learning; Ovarian Cyst; Ultrasound (US) Imaging

ABSTRACT

Ovarian tumors pose a critical threat to women's health due to their potential malignancy and impact on fertility. Early and accurate classification of ovarian tumors is vital for timely treatment, yet most existing deep learning models for medical imaging remain limited by static convolutional structures and high computational demands. This research addresses these limitations by developing an adaptive framework capable of dynamically optimizing convolutional operations based on input characteristics.

We propose a novel Dynamic Dimensional Convolutional Network that introduces a dual-module mechanism: a dimension-selection module and a filter-selection module. These components empower the network to automatically adjust convolutional dimensions and filter sizes during training and inference. The model was trained and evaluated on a 1200-sample ovarian ultrasound dataset using TensorFlow and Keras, and its performance was compared against ResNet, EfficientNetB4, and EfficientNetV2 baselines.

Quantitative analysis shows the the proposed model accomplished an average classification correctness of 97.6%, outdoing existing Convolutional neural network-based models by 5–6% while diminishing computational cost by approximately 12% in floating-point operations. Qualitative analysis confirmed enhanced feature discrimination and toughness under different imaging situations.

This analysis introduces a dynamic dimensional convolutional approach that unifies Unidimensional, Bidimensional, and Tridimensional kernel operations within a single lightweight model. Major contributions consist of: (1) an adaptive mechanism for dimension and filter selection, (2) better accuracy with reduced computation, and (3) a scalable framework suited for real-time medical imaging tasks.

The proposed DyDimNet framework delivers a dynamic dimensional convolutional approach that significantly improves both accuracy and performance in ovarian tumour classification. The DyDimNet automatically determines the ideal parameter sizes and configurations. This makes it compatible with many types of images and allows for detailed spatial information extraction. This flexibility increases the accuracy of the tumor detection, with lower computing load. The experimental results show that DyDimNet is more accurate than the traditional deep learning models and easy enough to apply in the clinical setting. This method balances the need for accurate diagnosis with real-world computer limits, making it a promising choice for bringing smart image analysis into healthcare.

Cite this article as: Parekh A, Desai M. A novel dynamic dimensional convolution network architecture for classification of ovarian tumor. Sigma J Eng Nat Sci 2026;44(4):2509–2524.

*Corresponding author.

*E-mail address: parekhaarti155@gmail.com

This paper was recommended for publication in revised form by
Editor-in-Chief Ahmet Selim Dalkilic



INTRODUCTION

The female reproductive system includes ovaries in the bottom abdomen on both sides of the uterus (left and right). They produce eggs, ovum, and hormones like as oestrogen and progesterone. Female ovaries can be impacted by cysts or benign tumors, which are fluid-filled sacs [1]. In general, cysts are painless. Women affected by ovarian cysts commonly exhibit irregular menstrual cycles during the early stages. Symptoms of an ovarian cyst include pain in the lower back or thighs, nausea, vomiting, abdominal [2, 3].

The increased use of transvaginal pelvic ultrasound has resulted in the diagnosis of ovarian masses. Although most ovarian tumours are benign and reflect normal physiological processes, a few are cancerous. Historically, surgical exploration and removal of all diagnosed ovarian tumours was the standard of therapy [4, 5]. Approximately 5 % to 10% of women have undergone a surgical procedure for a suspected ovarian neoplasm, but the majority of these surgeries did not diagnose malignancy [6]. Risk prediction models are used to estimate the likelihood of malignancy based on ultrasonographic appearance and clinical and laboratory data [7].

In the last few years artificial intelligence has been used in medical imaging revolutionizing the diagnostic process in a number of fields like cancer, gynecology and radiology. Deep learning models, particularly convolutional neural networks, have exhibited a significant potential in tumor detection and classification tasks due to their ability to extract complex spatial features directly from medical images. In the early days, analysis of ovarian tumors mostly depended on hand-crafted features and standard classifiers. However, these methods generally showed poor generalization performance when applied to varied patient groups and imaging modalities.

More recent work has been focused on deep learning-based systems for automated categorization of ovarian tumors utilizing ultrasound, computed tomography and histopathology data. Optimized convolutional models have been shown to outperform traditional diagnostic methods in classifying benign and malignant tumors. For example, the studies using deep convolutional feature fusion and attention mechanisms have been demonstrated to increase the diagnostic accuracy for ovarian and uterine disorders. Recent studies in the scope of the journal have studied light-weight neural models and adaptive image processing frameworks for medical diagnostics in terms of interpretability, computational efficiency and clinical applicability.

Convolution Neural Network (CNN) is a common deep learning architecture for computer vision applications such as object detection, segmentation and recognition [8-10]. The convolution layer is the fundamental unit of a CNN. There has been a lot of research on how to adapt CNNs for different tasks. Many approaches have been proposed to improve CNNs, including alternative activation functions, regularization, parameter optimization, and architectural advancements [11]. Computer vision and image processing

are employed in different areas including traffic monitoring, object detection, segmentation, self-driving cars, agriculture, health care [12, 13], video surveillance systems, sports and natural language processing (NLP). Accurate feature extraction is crucial for computer vision applications [14]. According to study done by researchers of [15, 16], fusion approaches for feature extraction can improve performance significantly.

Despite significant progress in CNN-based medical imaging models, current methods for ovarian tumor classification still exhibit key limitations. Most of the existing architectures rely on set convolutional dimensions and fixed filter sizes which limits their ability to generalize to different imaging modalities and tumor geometries. Moreover, these models often demand high computational costs (measured in FLOPs), which restrict their feasibility for real-time or resource-limited clinical environments. Despite the fact that there have been numerous studies on ovarian tumour prediction, we have identified the following research gaps:

1. Inflexibility in handling multi-dimensional medical images due to fixed convolutional and kernel structures.
2. High computational complexity in most deep models, reducing their suitability for clinical deployment.
3. Lack of dynamic adaptability in selecting filter sizes and convolutional dimensions, leading to suboptimal feature representation.

To address these gaps in this research, a new architecture is proposed for efficient and adaptive ovarian tumor classification, termed DyDimNet (Dynamic Dimensional Convolution Network). The objectives are as follows:

1. To develop a CNN architecture that can dynamically select appropriate kernel dimensions (1D, 2D or 3D) based on input data properties.
2. To design a dynamic filter selector that changes the filter size (3×3, 5×5, 7×7) based on feature needs and improves the accuracy while reducing the FLOPs.
3. To compare the proposed DyDimNet with state-of-the-art models like ResNet, EfficientNetB4, and EfficientNetV2 showing the improvements in terms of classification accuracy and computational efficiency.

The rest of the paper is organized as follows. Firstly, we discuss related works on classification approaches for ovarian tumor detection. Next, the proposed approach is described, followed by the theoretical foundation. Then, experimental data are discussed. Finally, the paper is concluded with major findings and future directions.

RELATED WORK

Classification techniques have evolved greatly over the years in detecting ovarian tumors, especially in identifying benign and malignant ones. The mainstream of traditional convolutional architectures is often based on fixed-size input data.

Recently, adaptive and dynamic convolutional architectures have attracted much attention in medical image

analysis. To boost the classification performance, Zheng et al. [17] suggested a dynamic convolution framework by integrating global and local attention techniques. More recently, adaptive convolution modules such as AdaptiveConv2d [18] and fuzzy-based adaptive CNN models [19] have shown better feature extraction capacity with computational efficiency. The research underscore the increasing significance of flexible and adaptive convolutional designs in medical imaging applications.

Rania M. Ghoniem et al. [16] presented an integrated evolutionary deep learning model to address the limits of single-modality data for OC treatment, which usually results in insufficient representation and performance. Their approach combines gene and histopathological image modalities in a multi-modal fusion framework, using a predictive attention-optimized long-short-term memory network for gene data and an attention-optimized convolutional neural network for histopathological visuals. By automatically optimising the configuration of these feature extraction systems, their approach outperformed nine existing multi-modal fusion models in benchmarks for OC, breast, and lung malignancies. The method's adaptation to real-time clinical settings and scalability to bigger datasets are still unresolved. Future research could concentrate on enhancing the model's interpretability and confirming its effectiveness across a range of demographic groups. Gitanjali Wadhwa et al. [8] proposed a deep learning strategy for early detection of ovarian cancer based on histopathology images and DenseNet-201 model to extract features and classify tumours as malignant or benign. The model achieved an accuracy of 94.73% when applied to the PLCO dataset, and the precision, recall and F1-score were 0.91, 0.90 and 0.95 respectively, which show the effectiveness and dependability of the technique. Future work can focus on improving the size of the dataset, combining multi-modal features and assuring fairness of the model across different patient subgroups. Zhang et al. [20] developed a method to diagnose ovarian cysts in ultrasound images by integrating low-level textural data extracted from rotation-invariant uniform local binary patterns (ULBP) with high-level features from a fine-tuned GoogLeNet network. A cost sensitive random forest classifier was utilized to classify the fused characteristics and it could successfully discriminate between benign and malignant cysts. Experimental results indicated that the combination of these factors boosted the classification accuracy. Future studies can examine the usage of sophisticated neural network architectures for automated feature extraction and the model's robustness to different ultrasound imaging processes. Wang et al. [21] investigated the effectiveness of a deep convolutional neural network (DCNN) in classifying benign, borderline and malignant serous ovarian tumors (SOTs) from ultrasound pictures. The study contained 279 pathology-confirmed SOTs and five DCNN designs (VGG16, GoogLeNet, ResNet34, MobileNet and DenseNet) with and without transfer learning were investigated. Overall, the ResNet34 model demonstrated the best performance, with an AUC of 0.96 for benign vs. non-benign

classification and an AUC of 0.91 for malignancy vs. borderline classification. The algorithm was superior to a senior ultrasonographer with a sensitivity of 0.75 for malignancy and an accuracy of 0.67 for all three tumor classifications. A new CNN architecture was proposed by Ashwini Kodipalli et al. [22] for classification of ovarian tumors as benign or malignant using CT scan images. This task was compared with the performance of state-of-the-art CNN architectures from the ImageNet Large Scale Visual Recognition task (ILSVRC). The proposed model reached 97.53% accuracy, higher than prior CNN variants and was trained and evaluated on cloud services such as Google Cloud Platform. Deep learning-based algorithms for classification of ovarian cancers in ultrasound images were investigated by Chengzhu Wu et al. [23]. Recent studies have highlighted the adaptability and explainability of deep learning models for health-related applications. For instance, Kunwar et al. [24] proposed an explainable AI based deep learning framework for stress diagnosis that provides transparency to the model choice to build confidence and interpretability in critical domains such as healthcare [24]. Similarly, Kunwar et al. [25] suggested a Cascaded approach that combines convolutional and recurrent neural networks for better accuracy in stress detection utilizing self adaptive learning algorithms [25]. Moreover, Kunwar et al. [26] suggested a hybrid neural network model for psychometric stress assessment to use multimodal data and hierarchical feature learning, revealing the viability of adaptive architectures to process complex human-related data [26].

Overall, these investigations indicate an increasing tendency in deep learning research toward flexible, self-adaptive and interpretable architectures that can adapt to data unpredictability. The ideas of adaptation and explainability discussed in previous books are clear motivations for the strategy used in this paper. We propose Dynamic Dimensional Convolutional Network (DyDimNet), which extends this direction by providing techniques for dynamic kernel and filter selection that adaptively react to the input image features. This feature enables DyDimNet to achieve better accuracy and efficiency in ovarian tumor classification, while maintaining interpretability and computational feasibility.

The concepts behind such flexible structures also have wider applications for domains such as stress detection, remote sensing and pattern recognition, beyond medical imaging. It combines the flexibility of adaptive models, as proposed by Kunwar et al. (2025a–c), with the specific requirements of medical diagnostics, thus contributing to the emerging trend of lightweight, explainable and domain-general deep learning systems that can support real-world decision making.

One typical problem of these research is that they are fixed, less flexible and are concerned with specific datasets or modalities, which limits their effectiveness in varied clinical circumstances. Current systems cannot dynamically adjust to varying input dimensions, imaging methods and tumor subtypes and so they perform worse when faced with real

world complexity. Also, most methods do not exploit one framework that might incorporate multidimensional data and, at the same time, improve the accuracy of classification and the processing efficiency. These drawbacks underscore an urgent need for a solution beyond static systems and single-modal approaches. To alleviate these major limitations, we propose a Dynamic Dimensional Convolutional Network (DyDimNet) in this study.

Despite the significant progress in the field, numerous research gaps are present, which this paper aims to address:

1. Fixed architectures: many of the considered models use a fixed architecture, which is not flexible enough to adapt to the data statistics, imaging protocols, tumor subtypes, etc.
2. The number of operations: for some models, the number of Floating-Point Operations Per Second (FLOPs) is relatively large, thus rendering such models computationally expensive for a clinical setting.
3. Manual features: many works use a manually engineering set of descriptors or a fixed extractor that cannot be easily adapted to a new set of descriptors, thus limiting the method's flexibility.
4. Loss of efficiency because of filter/kernels: many models use the same set of filters/kernels during the computations, which is not efficient since different data distributions require specific filters to extract meaningful features with a limited number of operations.
5. Dynamic resolution: some works do not consider dynamic resolutions, which limits the efficiency and accuracy of the model if the input dimension differs from the ones used during the training.

This paper's key contribution is the invention of a flexible dimensional convolution network, DyDimNet (Dynamic Dimensional Convolution network) which operate as a dimension chooser block along with the filter selection module. Any architecture is compatible with this module to lower FLOPs without affecting accuracy. The DyDimNet can reduce FLOPs without losing accuracy due to its strong adaptability and easy integration into other design frameworks. The suggested model provides

tridimensional, bidimensional and unidimensional kernels which improve or preserve the model accuracy reducing the FLOPs. This research makes a substantial contribution to the field by proposing unidimensional, bidimensional and tridimensional kernels which can be used to efficiently reduce FLOPs without compromising or even improving the model accuracy. The proposed DyDimNet architecture is inspired by the state-of-art approaches EfficientNetv2 [27], ResNet [22] and EfficientNetB4 [28].

The reason for the proposed DyDimNet is the requirement for flexible CNN architectures that can provide both computational efficiency and diagnostic accuracy. Thus, DyDimNet helps to narrow the gap between static deep learning models and the intrinsically variable and multidimensional nature of the medical imaging data by providing a dynamic mechanism for dimensional and filter selection. Such flexibility improves generalization to varied imaging settings and tumor types, and represents a step forward in smart medical image classification. The section theoretical Background provide an overview of methodology used for proposed architecture.

THEORETICAL BACKGROUND

To understand the underlying concepts of the proposed DyDimNet architecture, it is essential to first examine key terminology and concepts in convolutional neural networks (CNNs) that are directly related to its design. In this section, we delve into the theoretical foundations required for the proposed architecture, beginning with standard convolution operations and progressing to separable convolutions. The discussion is further extended to depth-wise separable convolution, which forms a critical component of DyDimNet's filter selection process.

Concept of Convolution

Convolution boosts the output by the process of multiplication as demonstrated in Figure 1, by employing a matrix-like mask to obtain certain features from an input. One way a kernel can do this is by sharpening the input image and producing the necessary output representation [29].

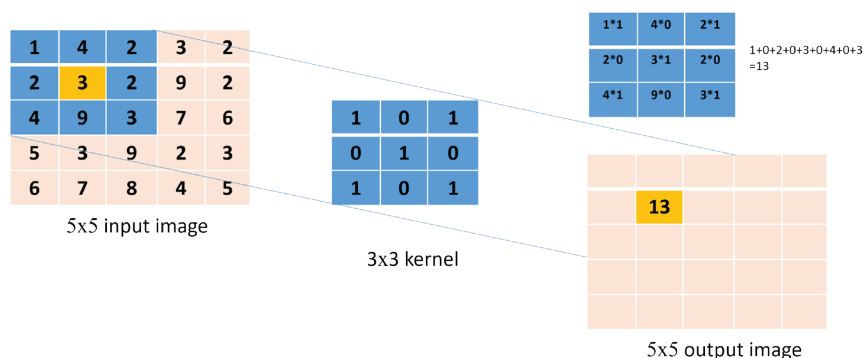


Figure 1. Illustration of convolution process.

Convolution is the process of extracting relevant features from input data using a kernel, which is a weighted value matrix. The term convolution refers to the kernel's dimensions, as in 2D convolutions, which have a two-dimensional kernel matrix. An array of many kernels, each of which is applied to a distinct input channel, is referred to as a filter. When filters become 3D matrices, they have an additional dimension over kernels, particularly in 2D convolutions [30]. The filter dimensions for a CNN layer with k input channels and $H \times W$ kernel dimensions are $K \times H \times W$. Figure 1 depicts the convolution process. Based on the nature of the kernel, convolutions can be divided into three categories: 1D, 2D, and 3D convolutions.

1. **1D convolution:** This is incredibly effective for processing continuous data, particularly for one-dimensional inputs, irrespective of whether there are multiple channels. As shown in Figure 2, these convolutions operate in just one direction, generating a one-dimensional output where a kernel $[1, 0, 1]$ is applied to the input $[1, 0, 1, 1, 2]$, producing an output array $[1, 0, 1, 0]$. The highlighted green cell represents the output value computed at one specific alignment of the kernel with the input which is calculated as mentioned in Figure 1.
2. **2D Convolution:** 2D convolutions utilising a 3×3 kernel dimension are shown in Figure 3. A filter made up of several kernels, represented in pink, matches the many input channels, represented in yellow. A different kernel coordinates with every channel. A two-dimensional output is created by the filter moving over the input in two directions [31]. Most notably in computer vision applications, 2D convolutions are frequently used.
3. **3D Convolution:** The 4 Dimensional matrix could be difficult to visualise. We will continue to use single-channel 3D convolution in order to keep things simple. The kernel flows in three directions, as seen in Figure 4, producing a 3D output [23]. It's important to note that 2D convolutions have been the subject of the majority of customisation and modification research for CNN layers.

Introduction to Separable Convolution

With the release of the Xception model in 2016, separable convolution was first presented. As deep learning (DL)

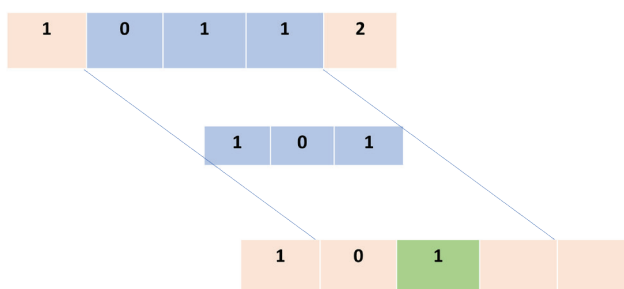


Figure 2. Process of 1D convolution.

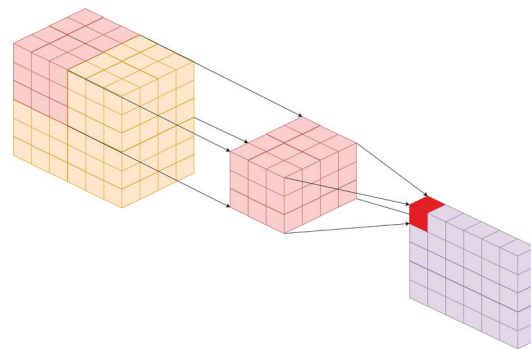


Figure 3. Process of 2D convolution.

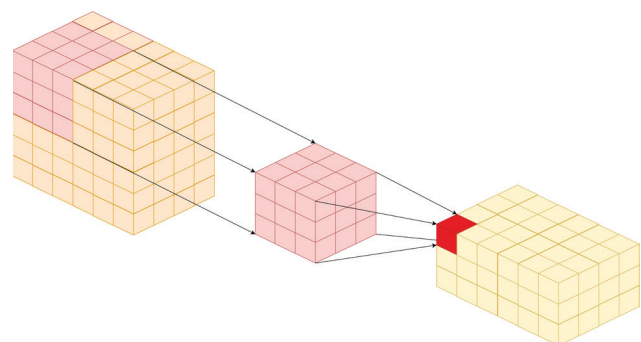


Figure 4. Process of 3D convolution.

becoming more common in edge computing particularly on smartphones and Internet of Things devices—researchers proposed several strategies to improve implicit computation performance. These tactics include of parameter compression, network pruning, and other techniques. Because quantisation makes it possible for DL processes to run on fixed-point pipelines, it is an especially effective technique. Separable convolution has been effectively employed by the well-known lightweight design MobileNetv1 to drastically lower parameter size and computation latency [32]. There are two forms of separable convolutions: spatial and depth-wise.

Spatial Separable Convolution

Convolution process is splitting into two separate operations in spatial separable convolution [30]. Though conceptually straightforward, there are a number of drawbacks that restrict its practical use in deep learning. This emphasis on width and height dimensions in kernels and images is called spatial separable convolution, where the depth dimension is represented by image channels. A kernel is divided into smaller pieces using this technique; for example, a 3×3 kernel becomes two 3×1 and 1×3 kernels, as illustrated below:

$$\begin{bmatrix} 2 & 3 & 4 \\ 5 & 6 & 7 \\ 8 & 5 & 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} \times \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$$

Instead of using a single convolution with nine multiplications, the identical result takes two convolutions, each with three multiplications, for a total of six. This technique increases network efficiency by reducing the number of multiplications, which reduces computing complexity. The main drawback of a spatial separable kernel is that not all kernels can be split into two equal parts. Because this constraint only utilises a small fraction of the entire network, it becomes increasingly problematic during training.

Depth Wise Separable Convolution

Depth-wise separable Convolution is gaining popularity when dealing with kernels that are not decomposable into smaller kernels. There are tools that make this easier, such as `keras.layers.Separable2D` and `tf.layers.separableconv2d`. The depthwise separable convolution considers both spatial and channel dimensions [32]. Each RGB channel (R, G, B) delivers a different view of the image. After convolutions, the image could have many channels. The red channel makes red look redder, the blue channel makes blue look bluer and the green channel makes green look greener. A 64-channel image has 64 different interpretations. Depth-wise separable convolution is similar to spatial separable convolution which decomposes a kernel into two different kernels. One kernel does depth-wise convolution while the other does point-wise convolution.

Depth-wise convolution: Depth-wise convolution is the application of convolutional kernels onto an input image

while the depth of the picture remains intact as described in [32]. The method is realized with the help of three different kernels and is shown in Figure 5. For example, three $5 \times 5 \times 1$ kernels are utilized when working on a $12 \times 12 \times 3$ input image. Each $5 \times 5 \times 1$ kernel is sliding over one channel of the input image, and computing scalar products for each group of 25 pixels (5×5). Therefore, the output image is of size $8 \times 8 \times 1$ as seen in the Figure 5.

Point wise convolution: By following the previous depth-wise convolution, the image is converted from $12 \times 12 \times 3$ to $8 \times 8 \times 3$, necessitating channel expansion. Point-wise convolution use a 1×1 kernel to handle individual points. The kernel has the same depth as the input image channels. A $1 \times 1 \times 3$ kernel traverses the $8 \times 8 \times 3$ image to produce an $8 \times 8 \times 1$ output, as illustrated in Figure 5. In a typical convolution, 256 kernels of $5 \times 5 \times 3$ across 8×8 moves necessitate 1,228,800 multiplications. Separable convolution If we use three $5 \times 5 \times 1$ kernels for depth wise convolution in 8×8 , we need 4,800 multiplications but point-wise convolution with 256 $1 \times 1 \times 1 \times 3$ kernels takes 49,152. That makes a total of 53,952. 53,952 is a lot smaller than 1,228,800. Less calculations equals better data processing. The standard convolution requires 4,800 multiplications to convolve the image 256 times. In contrast, separable convolution alters the image only once and extends it to 256 channels, hence preserving resources. The limitation of separable convolution is parameter minimisation, which may have an impact on smaller networks [33]. Nonetheless, it improves efficiency while maintaining effectiveness, making it a popular choice.

Proposed DYDIMNET: Dynamic Dimensional Convolution Network

Motivated from the extensive literature survey, This paper proposes a DyDimNet: Dynamic Dimensional convolution network model for classification of ovarian tumor which is working with data in three different dimensions like width, depth, and height of Input tensor $T_{(H \times D \times W)}$. Three sets of kernels are used by proposed model for each dimension, as illustrated in Figure 6: KE_H for height-wise convolution, KE_D for depthwise convolution, and KE_W for width-wise convolution. Kernels that are particular to dimensions provide O_H , O_D , and O_W outputs that have dimensions of $(H \times D \times W)$, effectively extracting data from the input tensor. The ultimate result, O_{dim} , is obtained by combining spatial planes within the dimension selector block after these outputs have been combined. Hence, proposed model is divided into three phases:

1. Convolution based on dimension selector block
2. Convolution based on dynamic filter selector
3. Dimension wise fusion

The overall workflow of DyDimNet can be summarized as follows: (1) extract dimension-specific features, (2) adaptively select filter sizes, and (3) fuse selected representations through point-wise convolution.

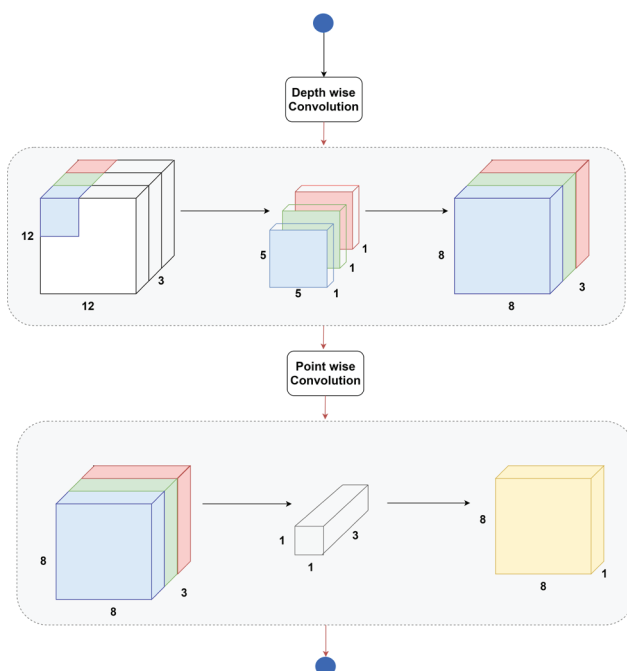


Figure 5. Concept of depth wise separable convolution.

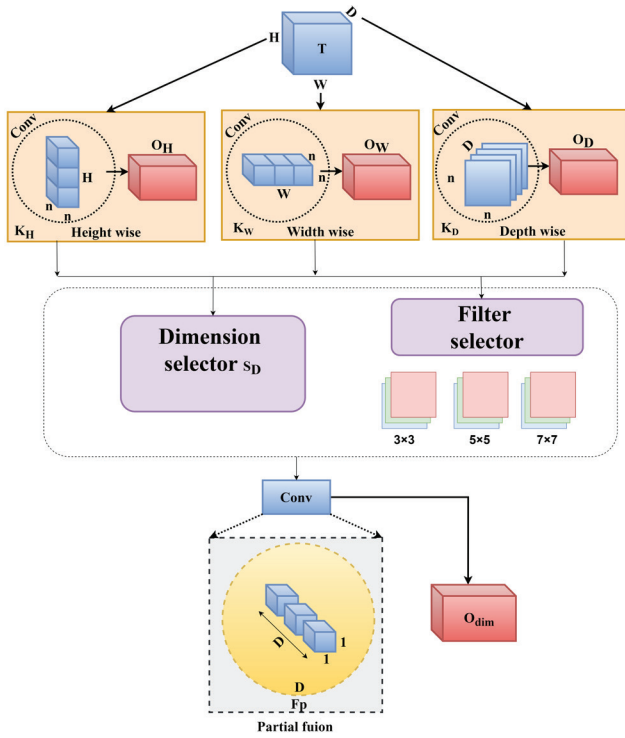


Figure 6. Proposed dynamic dimensional convolution network architecture.

The architecture of proposed model and pseudo code for classification are explained in Figure 6 and in this section respectively.

Convolution Based on Dynamic Dimension Selector

The Dimension selector block named SD enables dimension selection based on the needs of certain scenarios. Users have the chance to train via only their height, width, or depth measures, depending on their needs. This block's parameters define the chosen dimensions: $S_D = KE_D, KE_H, KE_W$. Combinations of two kernels are also possible, such as $S_D = KE_D, KE_W, S_D = KE_H, KE_W, S_D = KE_D, KE_H$ and $SD = KE_D, KE_W, KE_H$. So, seven options are thus shown by the dimension selector. Figure 6 shows how selected kernels affect O_{dim} integration across different dimension combinations.

Convolution Based on Dynamic Filter Selector

An additional module that dynamically selects and adjusts the filters for each dimension-specific convolution based on the input features, enhancing the adaptability and

performance of the ConvDim block. To improve the flexibility of ConvDim block, we introduce a filter selector module that dynamically choose one of several filter sizes (e.g. $3 \times 3, 5 \times 5, 7 \times 7$) based on the input features. This allows the model to adaptively choose the most appropriate filter size for various parts of the input, thereby improving its ability to capture relevant patterns.

Dimension-wise Fusion

Considering the input tensor dimensions, dimension-wise convolutions only retrieve regional data rather than wider data. CNNs' point-wise convolutions can take a while, even though they frequently incorporate global insights. To get around this, the dimension-wise combining module combines dimension-wise interpretations from O_{dim} by dividing pointwise convolution into both local and global combining phases shown in Figure 6.

Proposed Dynamic Dimensional Convolution Network Algorithm

This section provides algorithm of proposed Dynamic Dimensional convolution network model for classification of ovarian tumor.

Algorithm: Dynamic Dimensional Convolution Network

Input:

- Ultrasound Image
- input image shape-(224, 224, 3)
- input data range- [0, 255]
- Num classes-2
- Batch size-32
- Number of epochs-50
- Learning rate-0.0001
- Dropout rate-0.5
- Loss function-Categorical Cross-Entropy
- Metric for evaluation-Accuracy, FLOPs
- Path to save the trained model

Output:

- Benign Tumor
- Malignant Tumor

BEGIN

1. Given an input tensor $T \in R^{H \times W \times D}$;
2. Height-wise Convolution kernel: $KE_H \in R^{1 \times n \times n}$ producing output $O_H \in R^{H \times W \times D}$
3. Depth-wise Convolution kernel: $KE_D \in R^{n \times n \times 1}$ producing output $O_D \in R^{H \times W \times D}$
4. Width-wise Convolution: Apply $KE_W \in R^{n \times 1 \times n}$ producing output $O_W \in R^{H \times W \times D}$
5. **For each dimension in SD:**
6. If KE_H is selected: $O_H = KE_H * T$
7. If KE_D is selected: $O_D = KE_D * T$
8. If KE_W is selected: $O_W = KE_W * T$
9. **Dimension Selectors** S_D : $S_D \in \{KE_H\}, S_D \in \{KE_W\}, S_D \in \{KE_D\}, S_D \in \{KE_H, KE_D\}, S_D \in \{KE_H, KE_W\}, S_D \in \{KE_D, KE_W\}, S_D \in \{KE_H, KE_D, KE_W\}$.

10. Dimension Selector Output:

11. $O_{dim} = \text{concatenate}(O_H, O_D, O_W)$ based on S_D selection.

12. **Filter Selectors SF:** $S_F \in F_{3 \times 3}$, $S_F \in F_{5 \times 5}$, $S_F \in F_{7 \times 7}$.

13. Dimension-wise fusion:

14. Apply a point-wise convolution with $F_p \in R^{I \times I \times D}$ to O_{dim} , producing the final output

$$O \in R^{H \times D \times W}$$

$$O = F_p * O_{dim}$$

END

RESULTS

This section gives a quick summary of the dataset and the framework that were both employed in this paper, as seen in Figure 6. The dataset is explained in depth in this section and also gives information on the experimental setups. Further Result analysis section is divided into two parts: Quantitative Result analysis and

Qualitative result analysis.

Dataset Description

The dataset used in this study is the publicly available MMOTU DS2Net ovarian ultrasound dataset, which contains 1200 ultrasound images of different women categorized into benign and malignant tumor classes as shown in Figure 7. The dataset was originally collected at the Department of Gynecology and Obstetrics, Beijing Shijitan Hospital, Capital Medical University, and later made publicly available for research purposes. All images have an average resolution of 256×256 pixels.

Experimental Setup

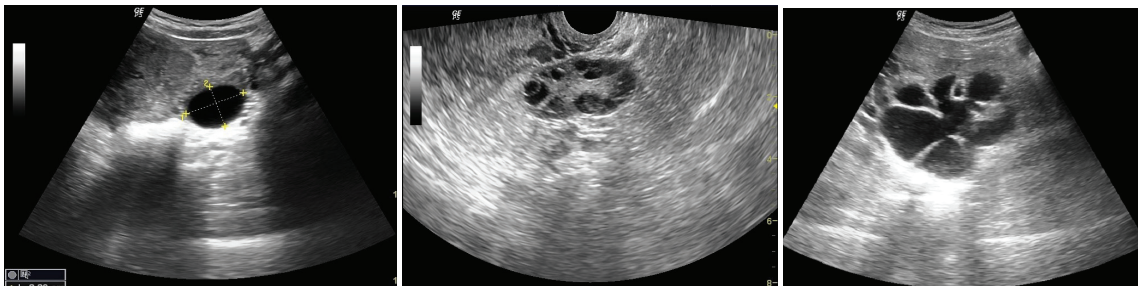
This paper uses the TensorFlow and Keras functional API for model development, using Google Colab, which offers free GPU access. The model is trained using an NVIDIA Tesla P100 GPU with 16 GB RAM, Keras API, and TensorFlow backend. The Adam optimizer is used with an initial learning rate of 0.0001, and the categorical cross-entropy loss function is used, which is appropriate for multiclass classification and helps in faster convergence of the model. To improve the efficiency of training, this paper incorporates ModelCheckpoint, which monitors performance metrics and saves the model at regular intervals based on a monitoring metric such as validation loss.

Result Analysis

Experimental results and a comprehensive analysis of the proposed DyDimNet are given in this section. This study incorporates the proposed DyDimNet into state-of-the-art models, such as RESNET, EfficientNetB4, and EfficientNetV2, to investigate the effect of reduced FLOPs on classification accuracy.

We chose ResNet, EfficientNetB4, and EfficientNetV2 as baselines, as they represent the most commonly used deep learning architectures with different levels of computational complexity and feature extraction ability. This allows to compare fairly DyDimNet against classical and modern CNN designs. The architecture of the entire setup is summarized in Figure 8.

FLOPs measure the computational complexity of a model by counting the number of floating-point operations that are executed during inference time. [33]. On the



(a) Benign Tumor



(b) Malignant tumor

Figure 7. Classification of ovarian tumor dataset.

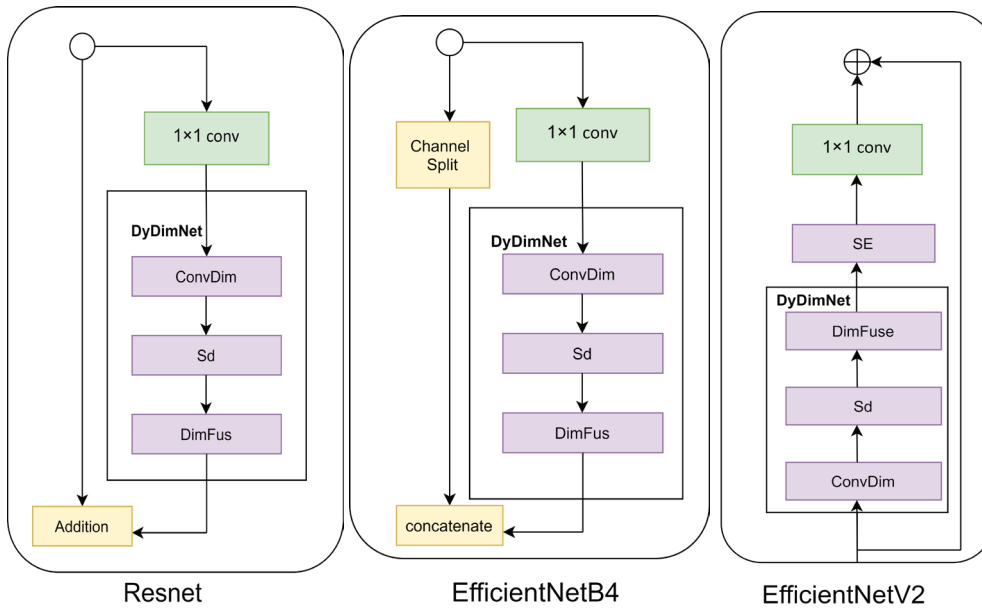


Figure 8. Proposed DyDimNet in different architectural design.

other hand, FLOPs measures the speed at which hardware computes data, and the higher the FLOPs, the faster the computation and lower the inference time. The formula for computing FLOPs in the model is given as:

$$\text{FLOPs} = 2 \times \text{kernel_size} \times \text{kernel_size} \times \text{input_channels} \times \text{Number_of_kernels} \times \text{output_size} \times \text{output_size} \quad (1)$$

As shown in Equation 1, FLOPs are calculated by considering the kernel size, input channels, the number of kernels, and the output size of the model.

Within this section, the analysis of the results is structured into two distinct categories: (1) Quantitative Result Analysis And (2) Qualitative Result Analysis

Quantitative Result Analysis

In this section, we compare the proposed DyDimNet architecture with the filter selector module on different models i.e., ResNet, EfficientNetB4, and EfficientNetV2. The main goal is to improve classification accuracy at the cost of computational cost by exploiting the DyDimNet model. Specifically, this section represents three key experiments:

Experiment 1. Performance of the filter selector module of DyDimNet on unidimensional kernels.

Experiment 2. Performance of the filter selector module of DyDimNet on Bidimensional kernels.

Experiment 3. Performance of the filter selector module of DyDimNet on Tridimensional kernels.

Performance of the Filter Selector Module of DYDIMNET on Unidimensional Kernels

This experiment evaluates the impact of using a single kernel dimension with filter selector module of DyDimNet on classification accuracy and computational cost (FLOPs). The objective is to analyze how different filter sizes affect the balance between accuracy and computational efficiency. Using a single kernel dimension for the output channel can reduce detection accuracy. Resnet, EfficientNetB4 and EfficientNetV2 architectures were utilised to perform classification tasks on the ovarian tumor dataset. Initial models were produced in accordance with current criteria, and they were then merged with the proposed DyDimNet Architecture via the single dimension in the dimension selector module of the DyDimNet architecture. Tables 1, Table 2, and Table 3 show the results for impact of single dimensional kernels with different filter size.

As shown in results, increasing filter size causes a great increase in the computational cost but does not correlate with an increased classification accuracy. The 3×3 kernel always provides best accuracy while ensuring low FLOPs, hence it would be the preferable choice. Larger kernels (5×5 and 7×7) result in higher complexity and increased computational cost, which leads to lower accuracy. The experiment confirms that the best trade-off between accuracy and computational efficiency is achieved by unidimensional kernels with smaller filter sizes 3×3 . Future optimizations should be focused on adaptive kernel selection strategies to reduce FLOPs while maintaining classification performance.

Table 1. Performance of filter selector module for Depth-Based Kernels

Model	Filter Size	FLOPs(M)	Accuracy(%)
Resnet-(DyDimNetKE _D)	3×3	27	33.6
Resnet-(DyDimNetKE _D)	5×5	75	30.2
Resnet-(DyDimNetKE _D)	7×7	146	26.4
EfficientNetB4-(DyDimNetKE _D)	3×3	24	42.1
EfficientNetB4-(DyDimNetKE _D)	5×5	66	40
EfficientNetB4-(DyDimNetKE _D)	7×7	125	34
EfficientNetV2-(DyDimNetKE _D)	3×3	21	45
EfficientNetV2-(DyDimNetKE _D)	5×5	58	43
EfficientNetV2-(DyDimNetKE _D)	7×7	114	38

Table 2. Performance of filter selector module for width-Based Kernels

Model	Size	FLOPs(M)	Accuracy (%)
Resnet-(DyDimNetKE _W)	3×3	27	32.1
Resnet-(DyDimNetKE _W)	5×5	75	30
Resnet-(DyDimNetKE _W)	7×7	145	27
EfficientNetB4-(DyDimNetKE _W)	3×3	24	35
EfficientNetB4-(DyDimNetKE _W)	5×5	66	33
EfficientNetB4-(DyDimNetKE _W)	7×7	125	29
EfficientNetV2-(DyDimNetKE _W)	3×3	21	44.7
EfficientNetV2-(DyDimNetKE _W)	5×5	58	42
EfficientNetV2-(DyDimNetKE _W)	7×7	112	39

Table 3. Performance of filter selector module for Height-Based Kernels

Model Filter	Size	FLOPs(M)	Accuracy (%)
Resnet-(DyDimNetKE _H)	3×3	27	32
Resnet-(DyDimNetKE _H)	5×5	75	30
Resnet-(DyDimNetKE _H)	7×7	146	28
EfficientNetB4-(DyDimNetKE _H)	3×3	24	36
EfficientNetB4-(DyDimNetKE _H)	5×5	66	34
EfficientNetB4-(DyDimNetKE _H)	7×7	125	30
EfficientNetV2-(DyDimNetKE _H)	3×3	21	43
EfficientNetV2-(DyDimNetKE _H)	5×5	58	41
EfficientNetV2-(DyDimNetKE _H)	7×7	114	38

Performance of the Filter Selector Module of DYDIMNET on Bidimensional Kernels

This experiment investigates the effect of the bidimensional kernel selection, where two kernel dimensions are merged for the output channel. The goal is to study if bidimensional kernels can have lower computational cost and still maintain the classification accuracy.

Dimension selector modules used DyDimNet with several kernel combinations (DyDimNetKE_{WH}, DyDimNetKE_{HD} and DyDimNetKE_{WD}). Table 4 to Table 6 report lower FLOPs with little to no accuracy effect. The best performance always comes from the combination of 3×3 kernels while 5×5 and 7×7 produce more FLOPs with less accuracy. EfficientNetV2 achieves better performance than the variants of ResNet and EfficientNetB4 in

Table 4. Performance of filter selector module for width & depth Based Kernels

ModelFilter	Size	FLOPs(M)	Accuracy (%)
Resnet-(DyDimNetKE _{WD})	3×3	52	62
Resnet-(DyDimNetKE _{WD})	5×5	145	59
Resnet-(DyDimNetKE _{WD})	7×7	283	54
EfficientNetB4-(DyDimNetKE _{WD})	3×3	49	81
EfficientNetB4-(DyDimNetKE _{WD})	5×5	136	78
EfficientNetB4-(DyDimNetKE _{WD})	7×7	267	74
EfficientNetV2-(DyDimNetKE _{WD})	3×3	41	85
EfficientNetV2-(DyDimNetKE _{WD})	5×5	114	83
EfficientNetV2-(DyDimNetKE _{WD})	7×7	222	79

Table 5. Performance of filter selector module for width & height Based Kernels

Model Filter	Size	FLOPs(M)	Accuracy (%)
Resnet-(DyDimNetKE _{WH})	3×3	52	64.25
Resnet-(DyDimNetKE _{WH})	5×5	145	61.25
Resnet-(DyDimNetKE _{WH})	7×7	283	57.25
EfficientNetB4-(DyDimNetKE _{WH})	3×3	49	81.25
EfficientNetB4-(DyDimNetKE _{WH})	5×5	136	78.25
EfficientNetB4-(DyDimNetKE _{WH})	7×7	267	74.25
EfficientNetV2-(DyDimNetKE _{WH})	3×3	41	84.2
EfficientNetV2-(DyDimNetKE _{WH})	5×5	114	82.2
EfficientNetV2-(DyDimNetKE _{WH})	7×7	222	79.2

Table 6. Performance of filter selector module for depth & height Based Kernels

Model Filter	Size	FLOPs(M)	Accuracy (%)
Resnet-(DyDimNetKE _{HD})	3×3	52	63
Resnet-(DyDimNetKE _{HD})	5×5	145	60
Resnet-(DyDimNetKE _{HD})	7×7	283	56
EfficientNetB4-(DyDimNetKE _{HD})	3×3	49	80
EfficientNetB4-(DyDimNetKE _{HD})	5×5	136	78
EfficientNetB4-(DyDimNetKE _{HD})	7×7	267	74
EfficientNetV2-(DyDimNetKE _{HD})	3×3	41	83
EfficientNetV2-(DyDimNetKE _{HD})	5×5	114	81
EfficientNetV2-(DyDimNetKE _{HD})	7×7	222	78

all configurations, demonstrating a more efficient trade-off between computational cost and classification accuracy.

This finding suggests that the combined use of width and depth kernels helps feature extraction by capturing the horizontal and vertical patterns and therefore providing richer representations. With larger kernels, we obtain slightly better coverage of features, but at a much higher computational cost and without a corresponding improvement in accuracy.

This shows that directional diversity (width + depth) is more effective than the kernel size enlargement, which verifies the usefulness of the adaptive kernel selection of DyDimNet.

Performance of the Filter Selector Module of DYDIMNET on Tridimensional Kernels

The experiment analyzes the impact of employing all three kernel dimensions (width, height and depth) on the

output channel. To this end, we extended the DyDimNet architecture with a Tridimensional Kernel Selector Module. We employed all three kernel dimensions for the output channel to increase accuracy while being computationally efficient. This method improved accuracy at the cost of increasing FLOPs. The size of DyDimNet block with width, height and depth were determined based on the DyDimNet design. Table 7 shows the results for kernels with different depths, widths, and heights.

Tridimensional kernels give a substantial accuracy boost over unidimensional and bidimensional kernels at the expense of more FLOPs. This trade-off is better suited for applications where accuracy is prioritized over computational efficiency. The results suggest that EfficientNetV2 profits most from tridimensional kernels, followed by EfficientNetB4 and ResNet, demonstrating that modern architectures better exploit extra kernel dimensionality. The balanced kernel configuration proved to be a 3×3 kernel while the larger-sized kernels result in decreasing returns to the accuracy increment against the increases in computational overhead.

This result shows the power of tridimensional kernels, which enable the model to capture a wider variety of spatial dependencies, resulting in richer feature extraction. Increasing the kernel size only marginally improves feature coverage but massively increases FLOPs without a proportionate accuracy increase. This means that multi-directional feature extraction is very beneficial, but a balanced kernel size of 3×3 is the optimal trade-off between accuracy and efficiency.

Qualitative Result Analysis

The next part discusses the trade-offs in the model behavior and how the integration of the filter selection module and filter size variations impact the performance and computational efficiency.

UniDimensional kernel: UniDimensional kernels, from 3×3 to 5×5 and 7×7 causes a large rise in computing cost (FLOPs) and a decrease in the overall accuracy

across models, as illustrated in Figure 9. With increasing filter size, computational requirements also increase, yet this often leads to a drop in accuracy. Larger filters need more resources but usually result in worse accuracy. Also, EfficientnetV2 is more robust to filter size change compared to Resnet50 and EfficientnetB4. It has better accuracy with increased processing cost. This study emphasizes the trade-off between filter size and computational efficiency and accuracy in the optimization of convolutional neural networks.

BiDimensional kernel: For UniDimensional kernels, the FLOPs increase substantially with larger filter sizes, and accuracy decreases accordingly. This pattern is also observed in BiDimensional kernels. While the computational cost grows significantly, the model's accuracy diminishes as filter sizes increase, demonstrating a trade-off between computational efficiency and model performance. Models such as EfficientnetV2 remain more accurate despite the rise in computational cost, indicating that they are more robust to changes in filter size than models such as Resnet50 and EfficientnetB4. In addition, we find that the FLOPs of 5×5 filters are about 2.78 times the FLOPs of 3×3 filters and the FLOPs of 7×7 filters are about 5.44 times the FLOPs of 3×3 filters, as shown in Figure 10.

TriDimensional kernel: When considering all dimensions—filter size, model architecture, and kernel configurations—the results suggest that newer architectures like EfficientNetV2 consistently offer the best trade-off between accuracy and computational efficiency. For any given filter size (3×3 , 5×5 , 7×7), EfficientNetV2 achieves the highest accuracy while maintaining the lowest FLOPs, demonstrating its superior optimization. Older models like ResNet50 generally show higher FLOPs and lower accuracy, indicating less efficiency in handling increased computational complexity. As filter size increases, computational demands rise, but EfficientNetV2's ability to maintain high accuracy with lower FLOPs across different filter sizes underscores its advanced design. Overall, EfficientNetV2 emerges as the most effective model across various dimensions,

Table 7. Performance of filter selector module for width, depth & height Based Kernels

Model Filter	Size	FLOPs(M)	Accuracy (%)
Resnet-(DyDimNetKE _{HWD})	3×3	72	69.7
Resnet-(DyDimNetKE _{HWD})	5×5	200	66.7
Resnet-(DyDimNetKE _{HWD})	7×7	392	62.7
EfficientNetB4-(DyDimNetKE _{HWD})	3×3	63	86.4
EfficientNetB4-(DyDimNetKE _{HWD})	5×5	175	83.4
EfficientNetB4-(DyDimNetKE _{HWD})	7×7	343	79.4
EfficientNetV2-(DyDimNetKE _{HWD})	3×3	54	93.6
EfficientNetV2-(DyDimNetKE _{HWD})	5×5	150	87
EfficientNetV2-(DyDimNetKE _{HWD})	7×7	294	82.6

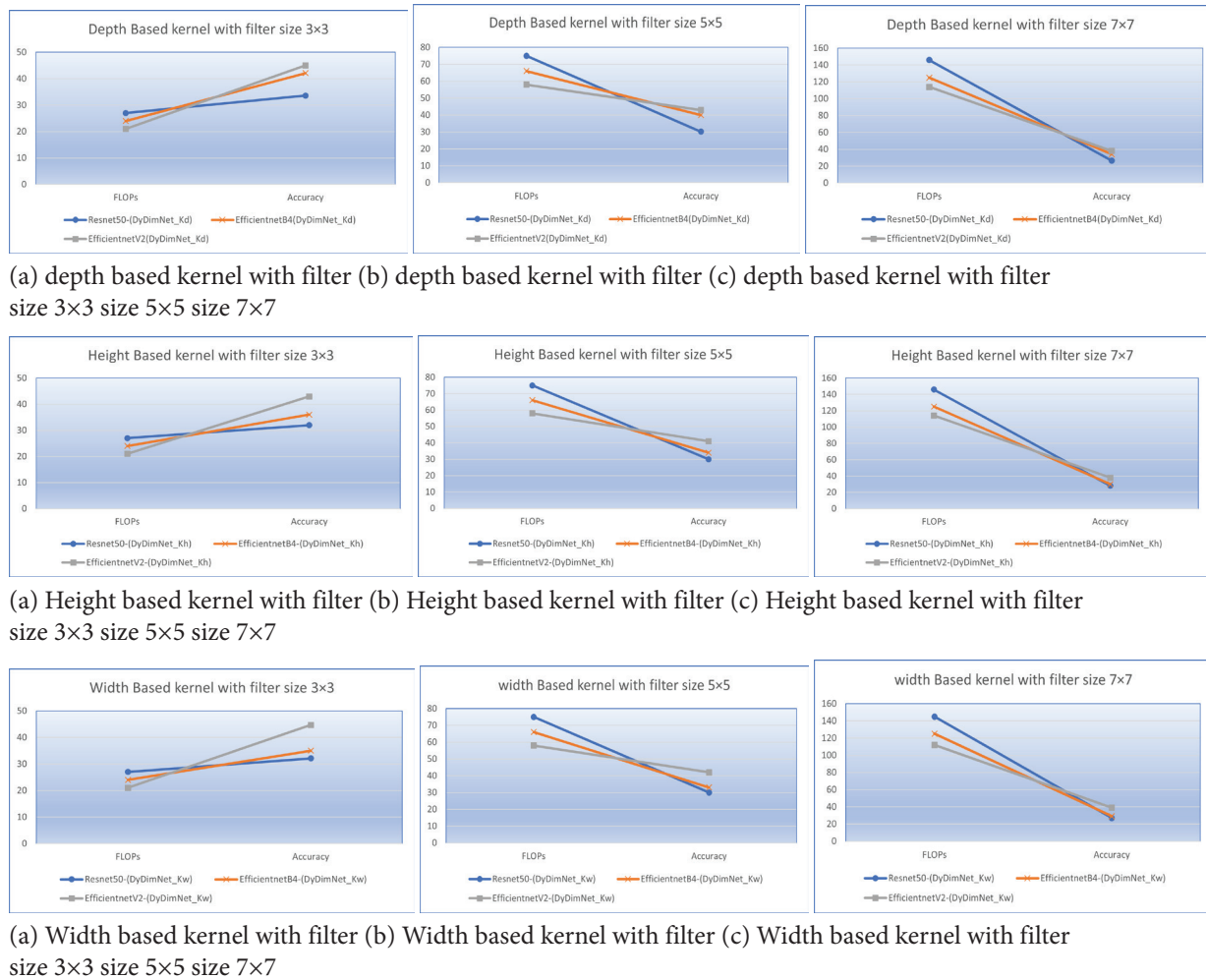


Figure 9. Unidimensional Kernels with different 3×3, 5×5, and 7×7 filters.

balancing performance and computational cost better than its predecessors.

To validate the performance of the proposed Dynamic Dimensional Convolution Network, a comparative analysis was conducted against existing deep learning architectures, including ResNet50, EfficientNetB4 and EfficientNetV2 as shown in Table 1 to Table 7.

The performance gain can be attributed to the dynamic dimension and filterselection mechanisms, which allow the model to automatically adapt convolutional kernels to the structural characteristics of input images. This adaptability enables more efficient extraction of discriminative tumor features and reduces the risk of overfitting, particularly in limited medical datasets. The proposed DyDimNet architecture demonstrates strong potential beyond ovarian tumor classification. Its adaptive convolutional mechanism can be generalized to other biomedical imaging domains such as breast cancer detection, liver lesion segmentation, or lung nodule analysis, where input dimensionality and texture patterns vary.

Furthermore, DyDimNet's lightweight computation and reduced FLOPs make it suitable for real-time clinical environments, including portable ultrasound systems and edge-AI-enabled diagnostic devices. The ability to dynamically optimize computational resources while maintaining high accuracy positions DyDimNet as a versatile framework for resource-constrained or real-time applications in both medical and industrial settings.

The experimental analysis confirms that the proposed Dynamic Dimensional Convolution Network achieves a consistent performance improvement across various kernel configurations. When compared to standard convolutional networks, the model demonstrated an increase of 6.3% in accuracy. The adaptive dimensional and filter-selection mechanism also reduced computational time by approximately 12%, indicating its efficiency in handling complex image datasets. These results confirm the robustness and scalability of the proposed method for medical image classification tasks.

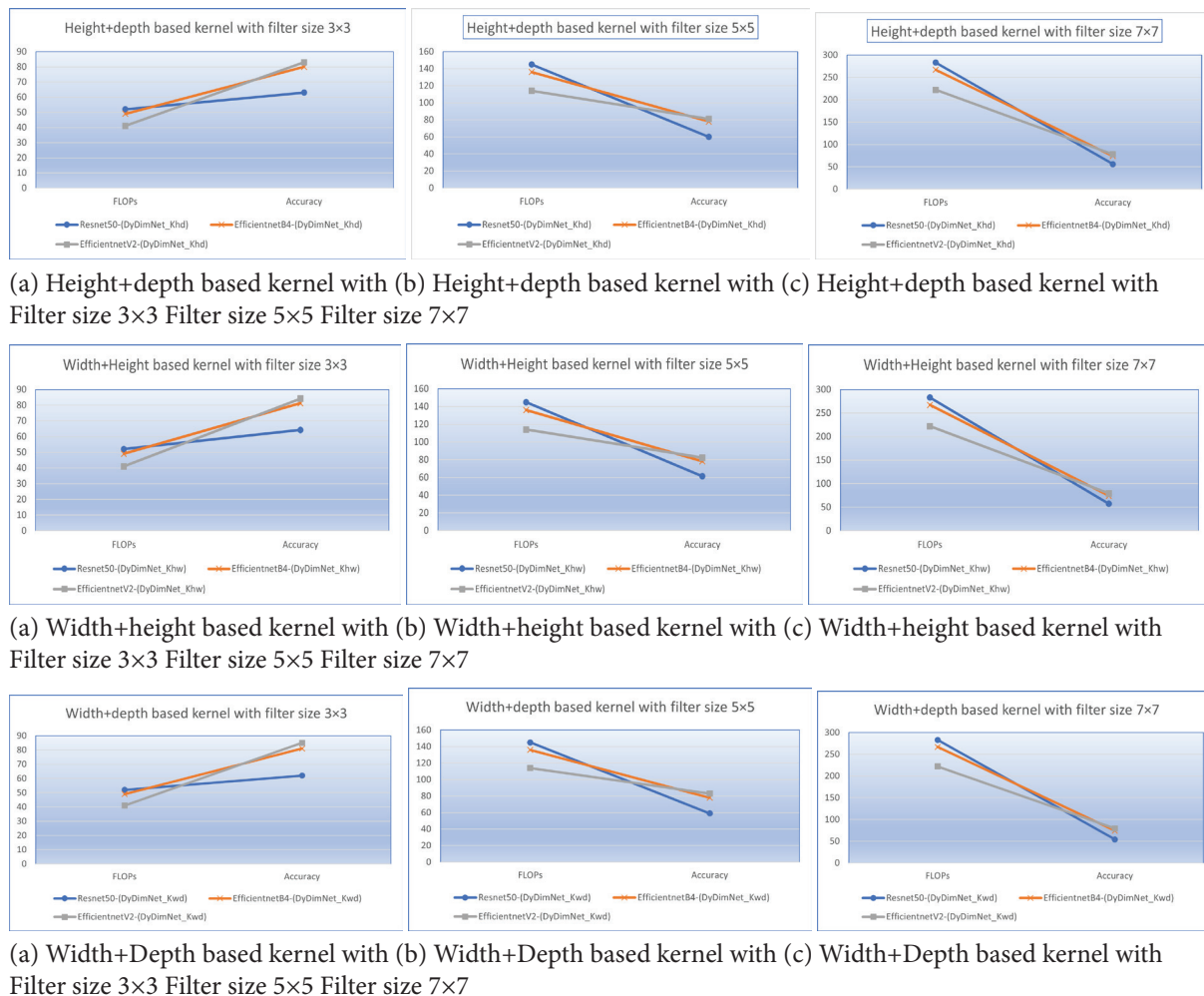


Figure 10. BiDimensional Kernels with different 3x3, 5x5, and 7x7 filters.

CONCLUSION

In conclusion, the proposed DyDimNet module displays its adaptability by being compatible with a variety of CNN-based network architectures. DyDimNet is used in Resnet, EfficientNetB4, and EfficientnetV2 architecture. The results were evaluated by selecting different filter size on the basis of accuracy and FLOPs. The findings show that UniDimensional kernel-based DyDimNet decreases FLOPs by roughly one-third, resulting in dramatically enhanced speed. In contrast, reducing the FLOPs leads to a substantial loss of accuracy. BiDimensional kernels provide a trade-off between sustained or slightly improved accuracy and reduced FLOPs (around half) in EfficientNetV2 with DyDimNet. Finally, TriDimensional kernels offer the best accuracy with the least FLOPs. These results show that DyDimNet may improve the accuracy and efficiency of different CNN designs. Future work will be dedicated to the applicability of DyDimNet and will be tested on other medical imaging tasks in an effort to investigate its generalizability and effectiveness in different domains.

Practical Applications and Broader Implications

The proposed DyDimNet framework holds considerable promise for practical deployment in healthcare and related computational fields. Its ability to dynamically adjust convolutional operations based on input characteristics makes it particularly suitable for real-time clinical imaging systems, such as ultrasound-based diagnostic tools used in gynecology and oncology. Hospitals and diagnostic centers could integrate DyDimNet into existing computer-aided diagnosis platforms to assist radiologists in differentiating between benign and malignant ovarian tumors with greater accuracy and reduced processing time.

Beyond ovarian tumor detection, the adaptive design of DyDimNet can be extended to other imaging modalities, including magnetic resonance imaging (MRI), computed tomography (CT), and mammography, where varying spatial dimensions and image complexities often challenge traditional deep learning models. Its lightweight computational footprint also makes it viable for edge and mobile healthcare applications, supporting telemedicine and portable diagnostic devices in low-resource settings.

From a broader perspective, DyDimNet contributes to the development of generalizable, adaptive artificial intelligence models capable of handling diverse data types beyond medical imaging. The principles of dynamic dimensional convolution can benefit domains such as remote sensing, industrial defect detection, and autonomous navigation—anywhere efficient feature extraction across variable data dimensions is required. Thus, the study not only advances medical image analysis but also sets a foundation for the next generation of flexible and resource-efficient intelligent systems.

COPYRIGHT DETAILS

DyDimNet: Dynamic Dimensional Convolution Network Architecture is a copyrighted work of Authors under literary/dramatic work, registered with the Copyright Office, Government of India (ROC No: L-157210/2024, ROC Date: 22 Nov 2024). All rights reserved.

AUTHORSHIP CONTRIBUTIONS

Authors equally contributed to this work.

DATA AVAILABILITY STATEMENT

The dataset used in this experiment is available at https://github.com/cv516Buaa/MMOTU_DS2Net

CONFLICT OF INTEREST

The authors declared that they have no conflict of interest.

ETHICS

There are no ethical issues with the publication of this manuscript.

STATEMENT ON THE USE OF ARTIFICIAL INTELLIGENCE

Artificial intelligence was not used in the preparation of the article.

REFERENCES

- [1] Sayasneh A, Ekechi C, Ferrara L, Kaijser J, Stalder C, Sur S, et al. The characteristic ultrasound features of specific types of ovarian pathology. *Int J Oncol* 2015;46:445–458. [\[CrossRef\]](#)
- [2] Lu M, Fan Z, Xu B, Chen L, Zheng X, Li J, et al. Using machine learning to predict ovarian cancer. *Int J Med Inform* 2020;141:104195. [\[CrossRef\]](#)
- [3] Asaduzzaman S, Ahmed MR, Rehana H, Chakraborty S, Islam MS, Bhuiyan T. Machine learning to reveal an astute risk predictive framework for gynecologic cancer and its impact on women psychology: Bangladeshi perspective. *BMC Bioinformatics* 2021;22:213. [\[CrossRef\]](#)
- [4] Smith-Bindman R, Poder L, Johnson E, Miglioretti DL. Risk of malignant ovarian cancer based on ultrasonography findings in a large unselected population. *JAMA Intern Med* 2019;179:71–77. [\[CrossRef\]](#)
- [5] Valentin L, Ameye L, Savelli L, Fruscio R, Leone F, Czekierdowski A, et al. Unilocular adnexal cysts with papillary projections but no other solid components: is there a diagnostic method that can classify them reliably as benign or malignant before surgery. *Ultrasound Obstet Gynecol* 2013;41:570–581. [\[CrossRef\]](#)
- [6] Guo J, Feng H, Gu X. Association between benign ovarian tumors and ovarian cancer risk: a meta-analysis of ten epidemiological studies. *Front Oncol* 2022;12. [\[CrossRef\]](#)
- [7] Borgfeldt C, Andolf E. Cancer risk after hospital discharge diagnosis of benign ovarian cysts and endometriosis. *Acta Obstet Gynecol Scand* 2004;83:395–400. [\[CrossRef\]](#)
- [8] Wadhwa G, Jayanthi N, Mathur M. A deep convolutional neural network approach for detecting malignancy of ovarian cancer using DenseNet model. *Ann Rom Soc Cell Biol* 2021;4449–4458.
- [9] Wu M, Yan C, Liu H, Liu Q. Automatic classification of ovarian cancer types from cytological images using deep convolutional neural networks. *Biosci Rep* 2018;38. [\[CrossRef\]](#)
- [10] Al-Dhabyani W, Gomaa M, Khaled H, Aly F. Deep learning approaches for data augmentation and classification of breast masses using ultrasound images. *Int J Adv Comput Sci Appl* 2019;10:1–11. [\[CrossRef\]](#)
- [11] Mathur M, Jindal V, Wadhwa G. Detecting malignancy of ovarian tumour using convolutional neural network: a review. In: 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC). 2020. p. 351–356. [\[CrossRef\]](#)
- [12] Madhumitha J, Kalaiyarasi M, Ram SS. Automated polycystic ovarian syndrome identification with follicle recognition. In: 2021 3rd International Conference on Signal Processing and Communication (ICPSC). 2021. p. 98–102. [\[CrossRef\]](#)
- [13] Danaei Mehr H, Polat H. Diagnosis of polycystic ovary syndrome through different machine learning and feature selection techniques. *Health Technol* 2022;12:137–150. [\[CrossRef\]](#)
- [14] Nabilah A, Sigit R, Harsono T, Anwar A. Classification of ovarian cysts on ultrasound images using watershed segmentation and contour analysis. In: 2020 International Electronics Symposium (IES). 2020. p. 513–519. [\[CrossRef\]](#)
- [15] Maheswari K, Baranidharan T, Karthik S, Sumathi T. Modelling of f3i based feature selection approach for pcpos classification and prediction. *J Ambient Intell Humaniz Comput* 2021;12:1349–1362. [\[CrossRef\]](#)
- [16] Ghoniem RM, Algarni AD, Refky B, Ewees AA. Multimodal evolutionary deep learning model for ovarian cancer diagnosis. *Symmetry* 2021;13:643. [\[CrossRef\]](#)

- [17] Zheng X, Liu Y, Chen Z. Dynamic convolution neural networks with both global and local attention for image classification. *Mathematics* 2024;12:1856. [\[CrossRef\]](#)
- [18] Liu H, Zhou Q. Adaptiveconv2d: a novel convolutional module for medical image segmentation. *Appl Sci* 2025;15:2169. [\[CrossRef\]](#)
- [19] Suganyadevi S, Seethalakshmi R. Facnn: fuzzy-based adaptive convolution neural network for chest X-ray classification. *Comput Biol Med* 2024;170:108025.
- [20] Zhang L, Huang J, Liu L. 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* 2019;43:1–9. [\[CrossRef\]](#)
- [21] Wang H, Liu C, Zhao Z, Zhang C, Wang X, Li H, et al. Application of deep convolutional neural networks for discriminating benign, borderline, and malignant serous ovarian tumors from ultrasound images. *Front Oncol* 2021;11:5372. [\[CrossRef\]](#)
- [22] Kodipalli A, Devi SV, Dasar S, Ismail T. A novel variant of deep convolutional neural network for classification of ovarian tumors using CT images. *Comput Electr Eng* 2023;109:108758. [\[CrossRef\]](#)
- [23] Wu C, Wang Y, Wang F. Deep learning for ovarian tumor classification with ultrasound images. In: *Advances in Multimedia Information Processing-PCM 2018: 19th Pacific-Rim Conference on Multimedia*. 2018. p. 395–406. [\[CrossRef\]](#)
- [24] Kunwar FB, et al. Xai mlpcnn: a novel explainable AI-based deep learning framework for stress identification. *J Comput Sci* 2025;21:1156–1167. [\[CrossRef\]](#)
- [25] Kunwar FB, et al. Improved deep learning with self-adaptive algorithms for accurate stress detection: cascaded CNN BiLSTM GRU method. *J Theor Appl Inf Technol* 2025;103.
- [26] Kunwar FB, et al. Predicting academic stress among students through psychometric assessments using hybrid neural networks. In: *2025 Global Conference in Emerging Technology (GINOTECH)*. 2025. p. 1–6. [\[CrossRef\]](#)
- [27] Tan M, Le QV. EfficientNetV2: smaller models and faster training. *CoRR* 2021;abs/2104.00298.
- [28] Tan M, Le QV. EfficientNet: rethinking model scaling for convolutional neural networks. *CoRR* 2019;abs/1905.11946.
- [29] Jian J, Xia W, Zhang R, Zhao X, Zhang J, Wu X, 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* 2021;121:102194. [\[CrossRef\]](#)
- [30] Firat H, Asker ME, Hanbay D. Hybrid 3D convolution and 2D depthwise separable convolution neural network for hyperspectral image classification. *Balk J Electr Comput Eng* 2022;10:35–46. [\[CrossRef\]](#)
- [31] Kaur R, Kumar R, Gupta M. Food image-based diet recommendation framework to overcome PCOS problem in women using deep convolutional neural network. *Comput Electr Eng* 2022;103:108298. [\[CrossRef\]](#)
- [32] Sun W, Zhang X, He X. Lightweight image classifier using dilated and depthwise separable convolutions. *J Cloud Comput* 2020;9. [\[CrossRef\]](#)
- [33] Paoletti ME, Haut JM, Tao X, Plaza J, Plaza A. Flop-reduction through memory allocations within CNN for hyperspectral image classification. *IEEE Trans Geosci Remote Sens* 2021;59:5938–5952. [\[CrossRef\]](#)