I.J. Intelligent Systems and Applications, **2025, 6, 94-104**
Published Online on December 8, 2025 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijisa.2025.06.07

Optimized Octave Convolution Network Model for Histopathological Image Classification

Binet Rose Devassy*
APJ Abdul Kalam Technological University, Thiruvananthapuram, Kerala, India
Department of Electronics and Communication Engineering, Sahrdaya College of Engineering and Technology, Thrissur, Kerala, India
E-mail: binetrosedevassy1@gmail.com
ORCID iD: https://orcid.org/0000-0002-1551-136X
*Corresponding author

Jobin K. Antony
Department of Electronics and Communication Engineering, Rajagiri School of Engineering and Technology, Kochi, Kerala, India
E-mail: jobinka@rajagiritech.edu.in
ORCID iD: https://orcid.org/0000-0001-6204-4365

Dominic Mathew
Department of Applied Electronics and Instrumentation, Rajagiri School of Engineering and Technology, Kochi, Kerala, India
E-mail: dominicmathew@rajagiritech.edu.in
ORCID iD: https://orcid.org/0000-0002-3212-4476

Received: 10 June 2025; Revised: 21 August 2025; Accepted: 28 September 2025; Published: 08 December 2025

Abstract: Accurate histopathological image classification plays a crucial role in cancer detection and diagnosis. In automated cancer detection methods, extraction of histological features of malignant and benign tissues is a challenging task. This paper presents a modified approach on octave convolution to extract high and low-frequency features which help to provide a comprehensive representation of histopathological images. Proposed octave convolution model is used to perform histopathological image classification using three different optimization strategies. Firstly, an optimal alpha value of 0.5 is used to give equal importance to both high-frequency and low-frequency feature maps. This balanced approach ensures that the model effectively considers critical high-frequency features as well as low-frequency features of cancerous tissues. Secondly, high-frequency and low-frequency feature maps are extracted and down sampled into half the spatial dimension size to reduce the computational cost compared to standard CNN. Thirdly, training and validation was conducted using ReLU, PReLU, LeakyReLU, ELU, GELU and Swish activation functions. From the experiment, it was concluded that PReLU is the best activation function for capturing intricate patterns inherent in cancer-related histopathological images. Combining all these optimization strategies, the proposed method proved to provide a classification accuracy of 93% and also to reduce the computational cost by 50%. Performance validation against pre-trained models, CNN variants and vision transformer-based models has also been conducted, which proved superior performance of the proposed model.

Index Terms: Histopathological Image, Octave Convolution, High and Low Frequency Feature Extraction, Spatial Down Sampling.

1. Introduction

Cancer stands as a significant global health concern, imposing a considerable burden of morbidity and mortality, ranking as the second most prevalent cause of death worldwide [1]. As per the study of International Agency for Research on Cancer (IARC), one in five individuals worldwide develops cancer during their lifetime. The COVID-19 pandemic has significantly disordered the diagnosis and treatment of cancer, resulting in a notable increase in advanced-stage cancer cases and mortality rates [2]. In this global scenario, proper diagnosis and treatment of cancer is one of the most significant

public health issues in our society. Biopsy is a critical procedure in cancer detection and analysis. It is known as the gold standard test for cancer. It allows the pathologists to examine the sample of suspicious tissue under a microscope. This process aids in confirming cancer presence, identifying its type and grade, determining its stage, and guiding treatment decisions. Biopsy is a crucial step in the cancer diagnostic and treatment process, providing valuable information that helps healthcare professionals deliver the most appropriate care for cancer patients.

Automated histopathological image analysis can provide objective and quantitative measurements of various features such as size, shape, and characteristics of nuclei, nuclei-cytoplasm features and tissue structures. These quantitative data can be valuable for disease diagnosis, treatment planning, tracking disease progression and research. Automated histopathological image analysis is not intended to replace the expertise of pathologists but rather to complement and enhance their capabilities. Pathologists remain essential for result validation, integration of clinical data, and to make critical diagnosis decisions based on the comprehensive assessment of patient cases.

Convolutional Neural Networks (CNNs) have been highly successful in various artificial intelligence-based tasks. But CNN has some limitations such as computation and memory overhead, loss of fine-grained details in pooling layers, and adaptation to information bottleneck. These limitations can be overcome by the octave convolution method. The octave convolution method is employed to overcome the constraints of conventional convolutional neural networks when dealing with multi-scale information and also to reduce computational overhead. It achieves this by processing feature maps at multiple frequencies, dividing them into high-frequency and low-frequency components. In conventional CNN, it processes the entire feature map at once, which can be computationally expensive and memory-intensive, especially for deeper networks and high-resolution images. Octave Convolution reduces the computation and memory overhead by processing feature maps at different frequencies separately. Pooling operations in CNNs reduce spatial resolution, leading to a loss of fine-grained details. Octave Convolution helps to contribute better feature extraction by processing the high-frequency band and low-frequency band separately. In very deep CNNs, the receptive field becomes larger, potentially causing an information bottleneck, where small-scale details may be neglected. Octave Convolution's design facilitates better information flow between frequency bands, helping to mitigate this bottleneck.

Deploying machine learning models, particularly those using histopathological images for cancer diagnosis, into clinical settings faces several challenges. AI models can offer real-time decision support, providing pathologists with immediate feedback or alternative diagnoses based on the histopathological images. This can be particularly useful in time-constrained environments where rapid and accurate decisions are needed.

The rest of the paper is structured as follows: Section 2 describes the related work relevant to the proposed method. Section 3 details the proposed model, which incorporates a modified octave convolution methodology. Section 4 explains the experimental results and performance analysis. Finally, Section 5 concludes the paper and outlines the future work.

2. Related Works

Automated Cancer detection using histopathological images contributes to the advancements of cancer research and accurate diagnosis. Handcrafted features designed by a domain expert are utilized in traditional machine learning algorithms and automatically learned discriminative features are utilized in deep learning models for image classification. Conventional feature extraction algorithms such as Scale Invariant Feature Transform [3], Extended Local Binary Pattern [4] and Gray Level Co-occurrence Matrix [5] are used for handcrafted feature extraction and for the analysis of histopathological images and tumor segmentation.

But it requires domain specific knowledge of an expert to identify the useful features. The supervised method of classification results in high computational load and low efficiency. Deep learning models, mostly Convolutional Neural Networks (CNNs) use convolutional layers to extract high level features from images. A hybrid CNN approach combines the possibilities of known hand-crafted features and autogenerated features of deep learning network for providing better result in image segmentation and classification task [6].

In the transfer learning approach, a pre-trained network model is initially trained on a large image database, acquiring knowledge that can be leveraged for new and distinct classification tasks [7]. From the pre-trained models, classification layers can be removed and new classification layers relevant to the particular problem can be added. The weights gained by the pre-trained models during the training can be used as weight initializers for the new problem of categorization of histopathological images. The fine tuning of weights based on the target task helps to capture useful features faster and also to improve accuracy of the model. Transfer learning methods for histopathological image analysis are used with pre-trained models such as VGG-16 [8], INCEPTION V3 [7], ResNet -50, DenseNet -161 [9]. Even though transfer learning offers improved generalization and enhanced performance, it still faces challenges such as domain shift, dataset bias, and limitations in feature transferability [10].

ResHist- a feature learning model inspired from ResNet-50 performs in a better way for learning the discriminative features. ResHist is a 152 layered model. Transfer learning is employed for the first 130 layers, and the final 22 layers of the model undergo training from scratch specifically on breast cancer histopathological images [11]. This model demonstrates a commendable performance with an accuracy of 84.34% and F1-score of 90.49% in the binary classification of histopathological images. A slide-based histopathology analysis framework is done using HipoMap which employs a feature aggregation approach. A pre-trained model is used to identify clinically relevant morphological features from patches within the slides. HipoMap -a one dimensional class representation is generated from this feature

aggregation and it is utilized by machine learning algorithms for the analysis of histopathological images [12].

Ensemble transfer learning enhances overall performance by combining individually trained pre-trained networks. Ensembling of the pre-trained networks which are fine tuned to the target task provides more robust results compared to the individual models [13]. In the multimodal lightweight XGBoost algorithm used for the prediction of brain tumor grading, an ensemble of XGBoost classifiers is used. Based on the performance, each classifier is weighted and their weighted sum leads to final prediction [14]. But ensemble models are computationally expensive by requiring more training time, more storage requirements with their associated parameters. To reduce the model complexity, squeeze and excitation network (SENet) model is employed. SENet plays a crucial role by recalibrating each feature map dynamically, enhancing its discriminative power through the selective amplification or attenuation of individual channels. This process relies on global spatial information, contributing to improved feature representation [15]. The SENet model with hybrid squeezing method involves dual modes of squeezing applied to feature maps. The first mode focuses on colour-based spatial squeezing, while the second involves channel-wise pooling [16]. This hybrid squeezing methodology aims to optimize the feature extraction process, combining both spatial and channel-level information for a more efficient and effective representation of complex patterns in the data. Based on the squeezed value, the feature maps are excited back. This will help to highlight or attenuate the corresponding feature maps based on the relevance of information content. This model shows a better performance of 91% accuracy and F1 score for binary classification of histopathological images with 35% reduction in the computational load. But the proposed method based on modified octave convolution helps to produce 93% accuracy with 50% reduction in computational cost.

3. Proposed Methodology: Modified Octave Convolution Method

Octave convolution is a novel approach in the realm of convolutional neural network layer which is designed to operate on two distinct frequency bands within the input feature map, dividing it into low-frequency band and high-frequency band. The primary goal of this technique is to enhance the network's ability to capture information at different scales. The low-frequency band is responsible for capturing broader contextual information, enabling the model to understand larger patterns and relationships in the data. On the other hand, the high-frequency band is dedicated to focusing on finer details and ensuring that the network can discern and emphasize intricate features.

Octave Convolution involves processing low-frequency and high-frequency feature maps separately through parallel sets of convolutional filters, enabling the network to selectively capture fine details in the high-frequency band and broader context in the low-frequency band. This technique enhances the model's ability to discern both fine-grained and global information within the input data. This enables the network to leverage the strengths of both types of filters and represent the multi-scale information present in the input feature map in a better way. Both high and low frequency components carry valuable information, and their significance depends on the specific analysis of the diagnostic task at hand.

For histopathological images, both high and low frequency features are important for the accurate diagnosis and analysis of cancer. High frequency features in histopathological images are indeed important for several reasons: These features capture fine details at a cellular level, such as the morphology of nuclei, cytoplasmic structures, cellular boundaries and the presence of abnormal tissue growth or invasion. At the same time, low frequency features capture the global patterns and structures in the histopathological images. Low frequency information represents the broader architectural organization of tissues, including the arrangement of cells, colour information, stroma, and extracellular components. These features provide contextual information that is crucial for understanding the overall tissue structure and identifying abnormalities within the larger histological context. Certain pathological conditions are characterized by alterations in the overall tissue structure rather than fine cellular details. For example, fibrosis, inflammation, or necrosis can lead to changes in tissue architecture that can be better captured by low frequency features. So, all these details are crucial for accurate identification and quantitative analysis of pathological conditions of tumor tissues.

In the conventional octave convolution model, low frequency feature maps are down-sampled or compressed to $0.5h$ x $0.5w$ size to reduce the spatial redundancy but the high frequency bands are kept at the same dimension [17]. In the proposed octave convolution model, both high-frequency and low-frequency feature maps are compressed to a size by half, aiming to alleviate redundancy in spatial information and reduce computational cost. To control the balance of distinct features between the high frequency and low frequency bands, octave convolution incorporates a ratio based mixing mechanism [18]. It involves a mixing parameter – alpha (α) that determines how much information from low frequency and high frequency feature maps should be mixed during the convolution operation. The concatenation mechanism in the modified octave convolution method combines the information from high-frequency and low-frequency bands. Concatenation ensures that the feature maps relevant to both fine-grained details and coarse features are stacked as the output of this custom layer. This helps in capturing more comprehensive and informative representations and contributes to the effectiveness of the Octave Convolution method in handling multi-scale information efficiently. Figure 1 illustrates the block diagram of the modified octave convolution model.

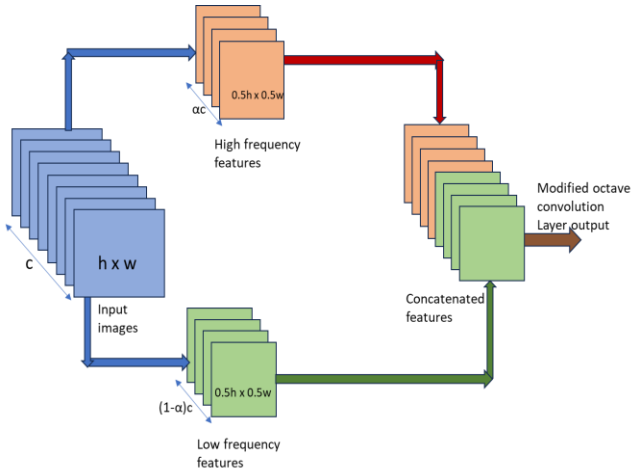


Fig.1. Block diagram of modified octave convolution model

Let input tensor and output tensor are denoted as X and Y respectively. Input tensor X is represented by $X \in R^{h \times w \times c}$ where h and w denote the spatial dimensions and c the number of feature maps or channels. This will be convolved with high frequency and low frequency kernel $W = [W^H, W^L]$. The input tensor will be factorized into X^H *and* X^L which corresponds to high frequency feature maps (X^H) and low frequency feature maps (X^L). Here, a ratio parameter α is selected to define the mixing ratio of number of channels for low frequency feature maps and to high frequency maps. The ratio α is selected to vary between 0 to 1 ($\alpha \in [0,1]$) to find the best balance of high and low frequency feature maps. Based on the ratio, α times low frequency feature maps and (1- α) times high frequency feature maps are assigned. These feature maps are reduced to half in spatial dimension using max pooling method.

The high frequency feature map with respect to location (p, q) is represented as convolution of input tensor with high frequency kernel W^H

$$Y^H{}_{(p,q)} = \sum_{i=1}^{H_w} \sum_{j=1}^{W_w} (X_{(p+i,q+j)} * W^H{}_{(i,j)}) \tag{1}$$

The low frequency feature map with respect to location (p, q) is represented as convolution of input tensor with low frequency kernel W^L is represented as

$$Y^L{}_{(p,q)} = \sum_{i=1}^{H_w} \sum_{j=1}^{W_w} (X_{(p+i,q+j)} * W^L{}_{(i,j)}) \tag{2}$$

These two frequency maps are max pooled. Therefore, spatial dimension of the feature map is reduced to half in size. After max-pooling, the high frequency feature map is having the dimension $X^H \in R^{\frac{h}{2} \times \frac{w}{2} \times (1-\alpha)c}$ and low frequency feature map is having the dimension $X^L \in R^{\frac{h}{2} \times \frac{w}{2} \times \alpha c}$. Then these pooled layers are concatenated to form the output tensor Y.

$$Y = f(pool\ (Y^H{}_{(p,q)}, 2)\ \ || f(pool\ (Y^L{}_{(p,q)}, 2) \tag{3}$$

In this modified octave convolution model, the number of floating-point operations (FLOPs) is reduced to 50%. The number of FLOPs in a Conv2D layer can be calculated as

$$FLOPs = 2 \times kernel\ size^2 \times input\ channels \times output\ channels \times \frac{input\ height \times input\ width}{stride^2} \tag{4}$$

For octave convolution layer, the number of FLOPs can be calculated as the summation of number of FLOPs of high frequency and low frequency feature maps *(hf_channel and lf_channel)*

$$FLOPs = FLOPs_{hf\_channel} + FLOPs_{lf\_channel} \tag{5}$$

By reducing the spatial size of both high-frequency and low-frequency feature maps into half, the number of floating-point operations is correspondingly reduced to 50%. It leads to reduction in the computational complexity of the model to 50%.

3.1. Proposed Model Architecture

Proposed model which utilizes a modified octave convolution model is shown below in Figure 2. Histopathological images of 128×128×3 size is applied to the proposed model. Firstly, they pass through a Conv2D layer of 128 filters to

extract the first level feature abstractions from the images. Then a series of modified OctConv2D_new layers are added to the model with 128, 64 and 32 filters. These layers are used with PReLU activation function and dropout rate of 20%. In these layers, high and low frequency features are extracted effectively. Followed by a flatten layer and dense layers of 256 and 128 nodes are implemented. Finally, a dense layer with a single node is included for binary classification. The final node is activated with sigmoid activation function to do the binary classification of benign and malignant tissues. The block diagram of the proposed model architecture using modified octave convolution layers is shown in Figure 2 and layer description with output shape is mentioned in Table 1.

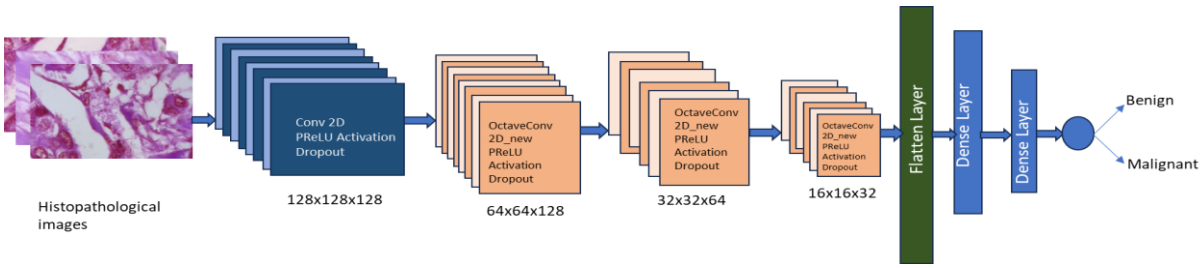


Fig.2. Proposed model architecture utilizing modified octave convolution network

The proposed architecture starts with an input layer which takes histopathological images from BreakHis dataset with size 128x128x3 is applied as input layer. It is followed by a Conv2D layer with 128 filters. These layers are responsible for learning local patterns and features from the input images. It is activated by the Parametric ReLU activation function. The Conv2D layer utilizes a set of adaptable filters on the input data to extract initial abstract features at the first level. It is followed by a modified octave convolution layer to capture both low frequency and high frequency features. The high frequency filters used in the octave convolution layer capture fine grained details whereas low frequency filters capture global information from the images. Three stages of the octave convolution layer used in the proposed model with 128, 64 and 32 filters for extracting multiple features from the different levels of abstraction. After extracting features from the convolutional layers, the Flatten layer is used to convert the multi-dimensional feature maps into a one-dimensional array vector. The flattened output is then passed to dense layers of 256 and 128 nodes. The final output layer is a single node layer with sigmoid activation function for classifying the histopathological images into binary classes of benign and malignant.

Table 1. Layers of the proposed model with output size

Layer	**Output Shape**
Input Layer	(None, 128,128, 3)
Conv2D Layer	(None, 128,128,128)
OctaveConv 2D_new_Layer	(None, 64, 64 ,128)
Activation (PReLU) Layer	(None, 64, 64 ,128)
Dropout Layer	(None, 64, 64 ,128)
OctaveConv 2D_new_Layer	(None, 32, 32, 64)
Activation (PReLU) Layer	(None, 32, 32, 64)
Dropout Layer	(None, 32, 32, 64)
OctaveConv 2D_new_Layer	(None, 16, 16, 32)
Activation (PReLU) Layer	(None, 16, 16, 32)
Dropout Layer	(None, 16, 16, 32)
Flatten Layer	(None, 8192)
Dense Layer	(None, 256)
Dense Layer	(None, 128)
Dense Layer (Sigmoid Activation)	(None, 1)

The model is trained using BreaKHis dataset with binary cross entropy loss function. The weights of the proposed model are optimized using Adam optimizer to minimize the loss function. Once trained, the model can be evaluated on test images in terms of performance measures such as accuracy, precision, recall and F1 Score by plotting the confusion matrix [19].

4. Results and Discussion

4.1. Dataset

The BreaKHis dataset comprises a total of 7,909 histopathological images extracted from breast tumor tissues obtained from 82 patients [20]. These biopsy tissue specimens were stained with hematoxylin and eosin dye and collected in collaboration with the P&D Laboratory – Pathological Anatomy and Cytopathology in Parana, Brazil [20]. The dataset is characterized by 2,480 benign samples and 5,429 malignant samples. For evaluation purposes, dataset is distributed for training, validation and testing in the ratio of 80:10:10. Therefore the distribution of images for training, validation and testing are 6327, 791 and 791 respectively.

4.2. Experimental Results for the Optimum Value of Alpha and Activation Function

Experiments were conducted to find an optimal value for hyper parameter alpha of octave convolution and also to find the best activation function for the proposed model. In the proposed method of modified octave convolution network, the appropriate proportion of high-frequency and low-frequency features from histopathological images is extracted and down sampled. The division of feature maps into these two groups is decided by the mixing factor - alpha (α). In the octave convolution process, a fraction α of the input feature channels is designated to extract low-frequency features while the remaining fraction ($1-\alpha$) of the feature maps is reserved for high-frequency features.

If α=0.3, 30% of the feature maps will be processed as low-frequency features, and 70% channels will be processed as high-frequency features. In case of first 'OctaveConv2D_new' Layer of the proposed model with 128 filters, α=0.3 produces 38 low frequency channels and 90 high frequency channels.

The proposed model is trained and validated for optimizing the α value (by varying α from 0 to 1) to effectively balance the extraction of fine-grained details and coarse patterns. This tuning process ensures that both types of features are processed efficiently, resulting in improved overall performance metrics. The model is evaluated with the BreaKHis dataset for 30 epochs. The tuning of the proposed model is done using the following hyperparameters to optimize model's performance. The nominal values selected are learning rate = 0.0001(1e-4), using the Adam optimizer, batch Size: 32, dropout rate: 0.5, applied after the dense layers to mitigate overfitting, weight decay (L2 Regularization): 0.0005, with early stopping implemented based on validation loss.

The training-validation accuracy and loss are plotted against the number of epochs as shown in Figure 3.a and 3.b.

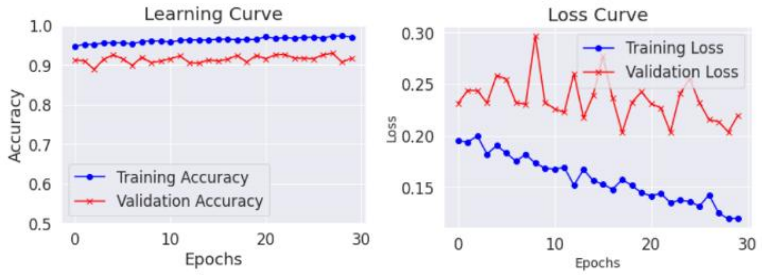


Fig.3. (a) Learning curve -training and validation accuracy (b)Loss curve -training and validation loss

The performance evaluation of the model is carried out over the test images. Out of the 791 test images, 530 images represent malignant samples and 261 images represent benign samples which results in an imbalanced class distribution. To ensure a fair performance evaluation, weighted averaging is applied. This approach prevents the model from being biased towards the majority class while accurately reflecting the contribution of each class based on its proportion in the dataset. Mathematical formula for weighted average performance metrics is given below.

$$Precision_{weighted} = \sum_{i=1}^{n} w_i \left(\frac{TP_i}{TP_i + FP_i}\right) \text{ where } w_i \text{ is the weight for class } i \quad (6)$$

$$Recall_{weighted} = \sum_{i=1}^{n} w_i \left(\frac{TP_i}{TP_i + FN_i}\right) \quad (7)$$

$$F1_{weighted} = \sum_{i=1}^{n} w_i \left(2 \times \frac{Precision_i \times Recall_i}{Precision_i + Recall_i}\right) \quad (8)$$

$$Accuracy_{weighted} = \frac{TP+TN}{TP+TN+FP+FN} \quad (9)$$

The weighted average performance measures for different alpha values are plotted in Figure 4 and listed in the Table 2.

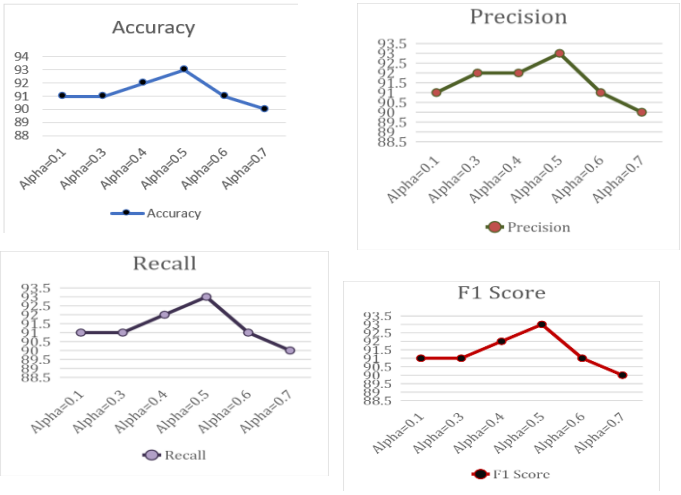


Fig.4. Performance metrics of the proposed model for different alpha values

Table 2. Performance measures of the proposed model for different alpha values

Alpha value	Accuracy	Precision	Recall	F1Score	Cohen's Kappa Score
α=0.1	91	91	91	91	80.5
α=0.3	91	92	91	91	80.7
α=0.4	92	92	92	92	81
α=0.5	93	93	93	93	82.5
α=0.6	91	91	91	91	81.7
α=0.7	90	90	90	90	81.5

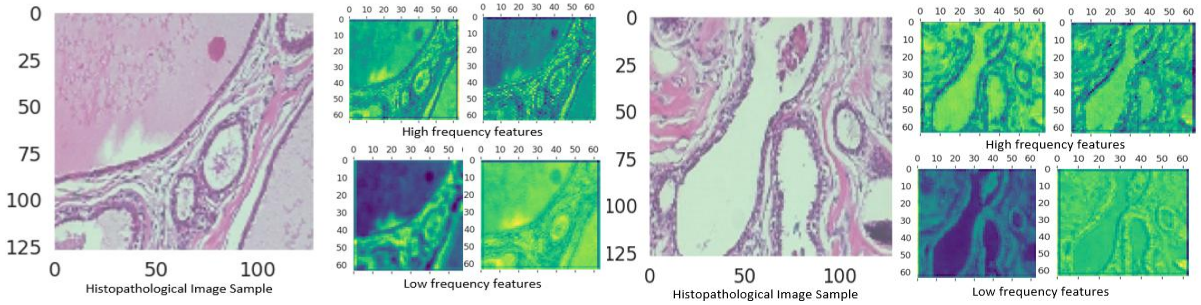


Fig.5. High frequency and low frequency features of OctaveConv 2D_new_Layer -1 for histopathological sample images

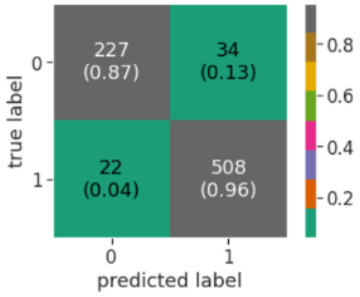


Fig.6. a. Confusion matrix

b. Calculated performance metrics

	precision	recall	f1-score	support
benign	0.91	0.87	0.89	261
malignant	0.94	0.96	0.95	530
accuracy			0.93	791
macro avg	0.92	0.91	0.92	791
weighted avg	0.93	0.93	0.93	791

The experiment shows that an alpha value of 0.5 gives the best performance for the model by providing the right balance between high frequency and low frequency features. This shows that the model is giving equal emphasis to the high frequency details and low frequency details. Low-frequency details such as arrangement of cells within a tissue, including their alignment, clustering, or distribution of high frequency or fine-grained details such as cell boundaries, textures and colour variation of nuclei are also informative for distinguishing benign and malignant tissues. Assigning equal weightage to high frequency and low frequency features helps to enhance the model's ability for the classification of cancer in histopathological images. High frequency and low frequency features of OctaveConv 2D_new_Layer -1 for

histopathological sample images are plotted in Figure 5. The model evaluation is illustrated using confusion matrix and calculated performance metrics values in Figure 6.a & b.

Another experiment was conducted to find the best activation function for the proposed model which improves the model's ability to capture features in histopathological images. The experiment showed that the Parametric Rectified Linear Unit (PReLU) activation function outperforms ReLU, Leaky ReLU, and ELU activation functions [21]. PReLU is a type of activation function that introduces an adjustable slope for negative values [22]. This adaptability in handling negative inputs contributes to its superior performance compared to other commonly used activation functions. Validation accuracy by progressing through different epochs is plotted in Figure 5. The findings underscore the effectiveness of PReLU in enhancing the learning capabilities of neural networks, making it a preferred choice in the classification of histopathological images for achieving improved model performance.

The PReLU function is described as follows.

$$f(x) = x\,; \qquad if\ x \geq 0;$$

$$f(x) = ax\,; \qquad if\ x \leq 0;$$

In ReLU activation, neurons with negative inputs always produce zero output, leading to vanishing gradient problem. PReLU addresses this issue by allowing a small negative slope which can help to prevent neurons from becoming inactive during training. The ability of PReLU activation function to model negative values ensures that the information in the dark regions of the images is also considered for the feature extraction stage. Dark nuclei are a valid symptom for malignancy in histopathological images. Therefore, PReLU activation function effectively captures features across high frequency and low frequency band and avoids vanishing gradient problem also. The graph shown in Figure 7 illustrates training accuracy and validation accuracy obtained over 30 epochs for different activation functions.

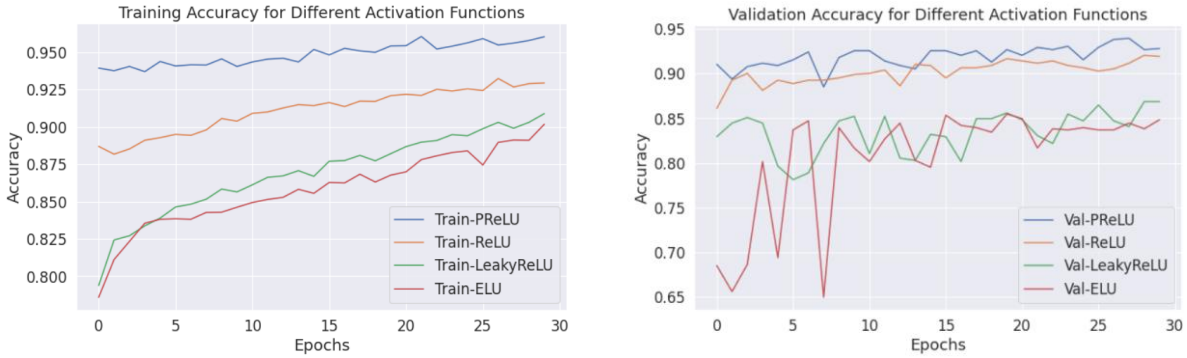


Fig.7. Training accuracy and Validation accuracy vs epochs for different activation functions

The proposed model is trained using various activation functions such as ReLU, PReLU, LeakyReLU, ELU, GELU and Swish to evaluate their performance in classifying benign and malignant cases. The performance evaluation metrics with different activation functions are obtained as shown in the Table 3.

Table 3. Performance metrics with different activation functions

Activation Function	Accuracy	Precision	Recall	F1 Score
ReLU	90	90	90	90
LeakyReLU	86	85	86	85
ELU	84	83	84	83
Swish	89	89	89	89
GELU	88	88	88	88
PReLU	**93**	**93**	**93**	**93**

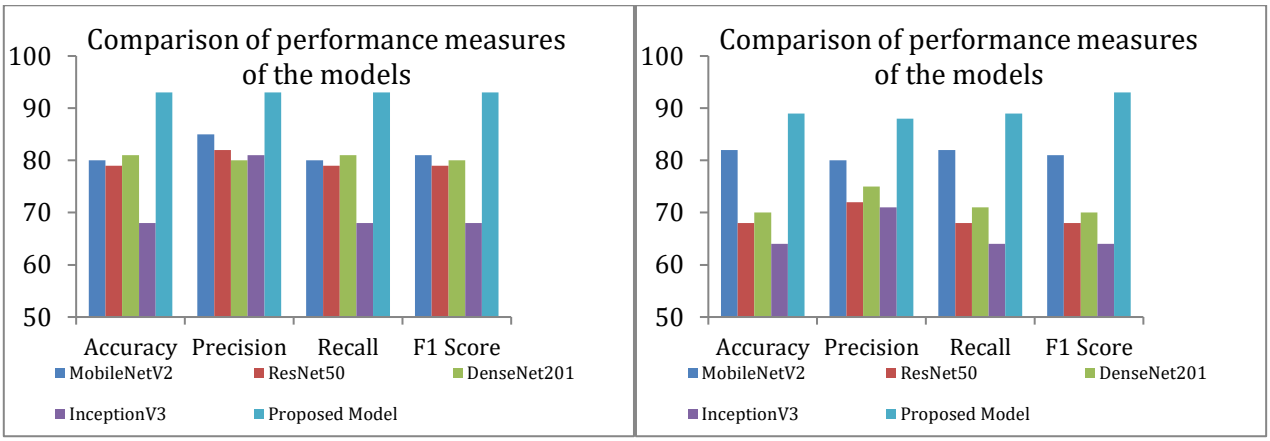
4.3. Comprehensive Performance Analysis of the Proposed Model Against Pretrained Models

The proposed model demonstrates better performance when compared to established pre-trained models, namely MobileNetV2, ResNet50, DenseNet 201, and InceptionV3. The evaluation involves disabling the top layers of these models and fine-tuning their weights on the BreakHis dataset. Despite the pre-trained models having deep-layered architectures and millions of learnable parameters, the proposed model surpasses them in terms of performance metrics such as accuracy, precision, recall and F1 score. The performance metrics is listed in Table 4. This indicates that the proposed model's design and training approach outperform the established pre-trained models in achieving enhanced

performance based on the evaluation metrics. Comparison of performance metrics of the pretrained models and the proposed model is plotted in the Figure 8.

Table 4. Comparison of performance metrics of the models

Model	Accuracy	Precision	Recall	F1 Score	Model	Accuracy	Precision	Recall	F1 Score
MobileNetV2	80	85	80	81	MobileNetV2	82	80	82	81
ResNet50	79	82	79	79	ResNet50	68	72	68	68
DenseNet201	81	80	81	80	DenseNet201	70	75	71	70
Inception V3	68	81	68	68	Inception V3	64	71	64	64
Proposed Model	**93**	**93**	**93**	**93**	**Proposed Model**	**89**	**88**	**89**	**88**



(a)Using BreakHis Dataset (b)Using BACH Dataset

Fig.8. Comparison of performance metrics of the models

4.4. Comprehensive Performance Analysis: Proposed Model vs. State-of-the-art Models

In addition to comparison with pretrained models, the performance of the proposed model was also evaluated against more recent architectures, such as the Vision Transformer (ViT) based model [23] and a CNN model utilizing vanilla convolution layers, configured with the same number of layers and filters as the proposed model. The performance of these models in terms of classification accuracy and computations are analysed.

Based on the experiments, the proposed modified octave convolution network achieved a classification accuracy of 93%, outperforming the ViT model, which achieved 90%, and the CNN model, which reached 89% on the same dataset for the same number of epochs. Additionally, the computational efficiency of the proposed model is compared in terms of number of trainable parameters and FLOPs of these models and it is listed in Table 5 [24]. Model evaluation parameters listed in the Table 5 demonstrated the superior performance of the proposed model in histopathological image classification. Histopathological images often contain both fine-grained cellular details and broader tissue patterns, which are challenging to capture simultaneously.

Table 5. Comparison of performance metrics of the models

Model	Accuracy	Number of trainable parameters	FLOPs ($G=10^9$)
Proposed model	93	2,058,576	1.42 G
CNN Model	89	8,665,313	5.61G
Vision transformer-based model	90	36,376,521	6.57G

The proposed modified octave convolution network, by separating high and low-frequency components, effectively captures both localized nuclear features and global features like cell arrangement and tissue invasions. In contrast, Vision Transformer-based models excel primarily in capturing global image representations due to their self-attention mechanism and may not be efficient in capturing fine grained details. The CNN models, though capable of capturing detailed local features, often fail to fully integrate broader contextual information like cell arrangement patterns. This makes them less efficient than the proposed model in handling the multi-scale nature of histopathological images. Also, the proposed model has a lower number of trainable parameters and requires fewer FLOPs, resulting in reduced model complexity and improved computational efficiency.

5. Conclusions and Future Work

Octave convolution is a method which processes the information in multiple frequency bands in order to capture both fine grained and coarse features. In the proposed method, the octave convolution model is modified by reducing the spatial dimension of high frequency and low frequency bands to half the size; which leads to 50% reduction in computational complexity. The proposed method of modified octave convolution model demonstrated an improvement in accuracy for histopathological image classification for distinguishing benign and malignant tissues by reducing computational complexity. The proposed model gives a promising performance measure with 93% test accuracy, precision, recall and F1 score respectively with a 50% reduction in computational load in terms of FLOPs. Also, the proposed model outperformed well with PReLU activation function compared to ReLU, Leaky ReLU and ELU. The performance of the proposed model is analysed with the pre-trained models MobileNetV2, ResNet50, DenseNet 201, and InceptionV3 and also evaluated against some recent architectures such as Vision transformer-based model and CNN Variants. From the performance measures, it is proved that the proposed model outperforms well against these state-of-the-art models in terms of accuracy and computational efficiency.

As the future work, model is to be validated against real-world dataset (clinical archives), which would better demonstrate the model's generalizability. Also, the model could be enhanced through multi-modal image fusion, combining histopathological images with MRI or PET scans to extract richer, more comprehensive information, improving diagnostic accuracy. Integrated into AI-driven diagnostic tools, it could assist pathologists in real-time clinical settings, streamlining workflows and enabling automated or semi-automated diagnostics. To address privacy concerns in medical data sharing, federated learning is proposed as a viable solution for training models across multiple institutions without centralized data storage [25]. This decentralized approach can enable collaborative medical research while safeguarding sensitive patient information. This is especially beneficial in resource-limited areas with fewer specialized pathologists, providing reliable diagnostics and scalable solutions. Future research could focus on optimizing the model for mobile devices or embedded systems, extending its reach to rural clinics and expanding its applicability in diverse healthcare environments.

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. “Global cancer statistics 2020: GLOBOCAN es timates of incidence and mortality worldwide for 36 cancers in 185 countries”. CA Cancer J Clin 2021;71:209-49.

[2] Tabatabaei, Z.; Wang, Y.; Colomer, A.; Oliver Moll, J.; Zhao, Z.; Naranjo, V. “WWFedCBMIR: World-Wide Federated Content-Based Medical Image Retrieval”, *Bioengineering* 2023, *10*, 1144. https://doi.org/10.3390/bioengineering10101144

[3] Lowe, D.G., “Distinctive Image Features from Scale-Invariant Keypoints”. International Journal of Computer Vision 60, 91–110 (2004). https://doi.org/10.1023/B:VISI.0000029664.99615.94

[4] F. Ullah, M. Nadeem, M. Abrar, "Revolutionizing Brain Tumor Segmentation in MRI with Dynamic Fusion of Handcrafted Features and Global Pathway-based Deep Learning," KSII Transactions on Internet and Information Systems, vol. 18, no. 1, pp. 105-125, 2024. DOI: 10.3837/tiis.2024.01.007.

[5] ŞabanÖztürk, BayramAkdemir, “Application of Feature Extraction and Classification Methods for Histopathological Image using GLCM, LBP, LBGLCM, GLRLM and SFTA”, Procedia Computer Science, Volume 132, 2018, Pages 40-46, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2018.05.057.

[6] Ullah F, Nadeem M, Abrar M, Al-Razgan M, Alfakih T, Amin F, Salam A. “Brain Tumor Segmentation from MRI Images Using Handcrafted Convolutional Neural Network. Diagnostics”. 2023; 13(16):2650. https://doi.org/10.3390/diagnostics13162650

[7] Ahmed S, Shaikh A, Alshahrani H, Alghamdi A, Alrizq M, Baber J, Bakhtyar M. “Transfer Learning Approach for Classification of Histopathology Whole Slide Images”, Sensors (Basel). 2021 Aug 9;21(16):5361. doi: 10.3390/s21165361. PMID: 34450802; PMCID: PMC8401188.

[8] Rakhlin, A., Shvets, A., Iglovikov, V., Kalinin, A.A. (2018), “Deep Convolutional Neural Networks for Breast Cancer Histology Image Analysis”. In: Campilho, A., Karray, F., ter Haar Romeny, B. (eds) Image Analysis and Recognition. ICIAR 2018. Lecture Notes in Computer Science, vol 10882. Springer, Cham. https://doi.org/10.1007/978-3-319-93000-8_83

[9] Muhammed Talo, “Automated classification of histopathology images using transfer learning”, Artificial Intelligence in Medicine, Volume 101, 2019, 101743, ISSN 0933-3657, https://doi.org/10.1016/j.artmed.2019.101743.

[10] R. W. Anwar, M. Abrar and F. Ullah, "Transfer Learning in Brain Tumor Classification: Challenges, Opportunities, and Future Prospects," 2023 14th International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2023, pp. 24-29, doi: 10.1109/ICTC58733.2023.10392830

[11] Kosaraju, S., Park, J., Lee, H. *et al.* “Deep learning-based framework for slide-based histopathological image analysis”. *Sci Rep* 12, 19075 (2022). https://doi.org/10.1038/s41598-022-23166-0

[12] Gour, Mahesh et al. “Residual learning based CNN for breast cancer histopathological image classification.” *Int. J. Imaging Syst. Technol.* 30 (2020): 621-635.

[13] Devassy, B.R., Antony, J.K. (2022). “Histopathological Image Classification Using Ensemble Transfer Learning”. In: Misra, R., Shyamasundar, R.K., Chaturvedi, A., Omer, R. (eds) Machine Learning and Big Data Analytics (Proceedings of International Conference on Machine Learning and Big Data Analytics (ICMLBDA) 2021). ICMLBDA 2021. Lecture Notes in Networks and Systems, vol 256. Springer, Cham. https://doi.org/10.1007/978-3-030-82469-3_18

[14] Ullah, Faizan & Nadeem, Muhammad & Abrar, Muhammad & Amin, Farhan & Salam, Abdu & Alabrah, Amerah & AlSalman, Hussain. (2023). “Evolutionary Model for Brain Cancer-Grading and Classification”. IEEE Access. PP. 1-1. 10.1109/ACCESS.2023.3330919.

[15] J. Hu, L. Shen and G. Sun, "Squeeze-and-Excitation Networks," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018, pp. 7132-7141, doi: 10.1109/CVPR.2018.00745.
[16] Devassy, B.R., Antony, J.K. "Histopathological image classification using CNN with squeeze and excitation networks based on hybrid squeezing". *SIViP* 17, 3613–3621 (2023). https://doi.org/10.1007/s11760-023-02587-y
[17] Y. Chen *et al*., "Drop an Octave: Reducing Spatial Redundancy in Convolutional Neural Networks With Octave Convolution," 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Korea (South), 2019, pp. 3434-3443, doi: 10.1109/ICCV.2019.00353.
[18] https://doi.org/10.48550/arXiv.2104.05345
[19] Dhar P, Garg VK, Rahman MA, " Enhanced Feature Extraction-based CNN Approach for Epileptic Seizure Detection from EEG Signals". J Healthc Eng. 2022 Mar 16;2022:3491828. doi: 10.1155/2022/3491828. Retraction in: J Healthc Eng. 2023 May 24;2023:9825862. PMID: 35340257; PMCID: PMC8942662.
[20] F. A. Spanhol, L. S. Oliveira, C. Petitjean and L. Heutte, "A Dataset for Breast Cancer Histopathological Image Classification," in IEEE Transactions on Biomedical Engineering, vol. 63, no. 7, pp. 1455-1462, July 2016, doi: 10.1109/TBME.2015.2496264
[21] Lukic V, Brüggen M. "Galaxy Classifications with Deep Learning". Proceedings of the International Astronomical Union. 2016;12(S325):217-220. doi:10.1017/S1743921316012771
[22] arXiv:1502.01852 https://doi.org/10.48550/arXiv.1502.01852
[23] Abimouloud, M.L., Bensid, K., Elleuch, M. *et al.* "Vision transformer based convolutional neural network for breast cancer histopathological images classification". *Multimed Tools Appl* (2024). https://doi.org/10.1007/s11042-024-19667-x
[24] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale". *https://arxiv.org/abs/2010.11929* .
[25] Ullah F, Nadeem M, Abrar M, Amin F, Salam A, Khan S. "Enhancing Brain Tumor Segmentation Accuracy through Scalable Federated Learning with Advanced Data Privacy and Security Measures". Mathematics. 2023;11(19):4189. https://doi.org/10.3390/math11194189

Authors' Profiles

Binet Rose Devassy is currently a research scholar at APJ Abdul Kalam Technological University, Kerala, India. She is working as Assistant Professor at Sahrdaya College of Engineering and Technology, Kerala, India. She obtained the Bachelor degree in Applied Electronics and Instrumentation from Government Engineering College, Calicut, Kerala, India and Master's degree in Applied Electronics from Anna University, Tamil Nadu, India. She has more than 20 years of teaching experience. Her areas of interest are in medical image processing, computer vision and artificial intelligence.

Dr. Jobin K. Antony: He received his M.Sc. in Electronics and M.Tech. degree from Cochin University of Science and Technology, Kochi, India. He completed his Ph.D. at the Indian Institute of Technology Madras. He also worked as a research scholar at Kyushu University, Japan, for a short period under the JASSO scholarship. He has more than 30 publications to his credit. He is currently working as a Professor and Dean in the Department of Electronics and Communication Engineering at Rajagiri School of Engineering & Technology, Kochi, India.

Prof. Dominic Mathew graduated from College of Engineering, Trivandrum, Kerala, India in Electrical Engineering and holds Post graduate degree in Power Electronics from CUSAT, Kerala. He has more than 23 years of industrial experience in the area of industrial automation and approximately 18 years of experience in academics. His areas of interest are in Industrial Automation, Signal Processing, Computer Vision and Artificial Intelligence. He has published around 34 papers related to these areas and is currently involved in teaching topics related to A.I. at the Rajagiri School of Engineering & Technology, Kakkanad, Kerala

How to cite this paper: Binet Rose Devassy, Jobin K. Antony, Dominic Mathew, "Optimized Octave Convolution Network Model for Histopathological Image Classification", International Journal of Intelligent Systems and Applications(IJISA), Vol.17, No.6, pp.94-104, 2025. DOI:10.5815/ijisa.2025.06.07