CancerSeg-XA: Enhanced Breast Cancer Histopathology Segmentation Using Xception Backbone with Attention Mechanisms

Alaa Mohamed Youssef, Wessam Hassan El Behaidy, and Aliaa Abdel-Haleim Abdel-Razik Youssif

Original scientific article

***Abstract*—Breast cancer remains a formidable health challenge requiring advanced computational tools for accurate diagnosis and treatment planning. This study hypothesizes that modifications to the DeepLabV3+ architecture, such as incorporating an attention layer and replacing the ResNet50 backbone with Xception, can significantly enhance segmentation accuracy and model stability for breast cancer histopathological images. To test this hypothesis, we evaluated the performance of the original DeepLabV3+ and three modified versions for semantic segmentation using the "Breast Cancer Semantic Segmentation" (BCSS) dataset, which provides pixel-wise annotations of breast cancer tissues. The proposed modifications include integrating an attention layer between the encoder and decoder (Model 1), replacing the ResNet50 backbone with an Xception backbone up to 'block5' (Model 2), and combining the Xception backbone with the attention layer (CancerSeg-XA). The models were implemented and trained in the Kaggle Notebook environment, and their performance was assessed based on training and validation accuracy. The results show that Model 1 improved the model stability and accuracy compared to DeepLabV3+, whereas Model 2 and CancerSeg-XA achieved significant accuracy improvements of 91.47% and 91.57%, respectively, over the baseline DeepLabV3+ accuracy of 85.7%. CancerSeg-XA demonstrated enhanced training stability, making it a promising approach for clinical application in breast cancer diagnosis and treatment.**

***Index terms*—Deep learning, Breast cancer segmentation, Histopathological images, ResNet50, DeepLabV3+, Xception backbone, Attention mechanism, BCSS.**

I. Introduction

Cancer remains one of the leading causes of mortality worldwide, with nearly 10 million deaths reported in 2020 [1].

It involves transforming normal cells into malfunctioning cells that rapidly multiply and invade other tissues. In 2022, breast cancer was the most common cancer among women globally [2, 3].

Many types of cancer can be treated successfully and even cured if detected early and managed appropriately. However, the process of diagnosing cancer is often time-consuming and requires pathologists to examine the stained tissue samples under a microscope meticulously. With the advent of digital scanners for whole-slide imaging (WSI), computer vision has emerged as a tool to significantly reduce the diagnosis time by automating the analysis of WSIs. WSI involves creating digital histopathological images through laboratory processes, which can be examined either manually under a microscope or digitized at varying magnifications using specialized scanners [4].

Histopathology image segmentation is a critical task in computer vision that involves precise delineation of structures and regions within WSIs. This process starts with high-resolution digital scanning of stained wax tissue slices on a glass slide at multiple magnifications, achieving a spatial resolution of 0.25 μm/pixel at 40X magnification. WSI images are stored in a multiscale pyramid resolution format, as illustrated in

Fig. 1 [5].

This provides a comprehensive view of the tissue specimen, enabling pathologists to meticulously perform the labor-intensive task of pixel-wise labeling of WSIs and annotating individual pixels corresponding to distinct tissue structures or pathological features [6]. Generating sufficient manually labeled data to train convolutional neural network (CNN) models poses a significant challenge because of the meticulous annotation required for gigapixel whole-slide images, a process that is both costly and time-consuming. This complexity is further amplified by the heterogeneity of malignant tumors because tumors of the same type can exhibit highly diverse morphological characteristics.

This intricacy necessitates expertise in clinical backgrounds for accurate pixel-level labeling, making it unsuitable for individuals lacking such qualifications [7]. Additionally, experts employ different magnification levels when annotating WSIs and adjust their focus to capture fine details and variations in the tissue morphology. Consequently, finding ways to alleviate the annotation burden remains an open problem, prompting the exploration of innovative approaches to reduce the efforts required for this crucial aspect of histopathological image segmentation [7].

The field of deep-learning-based medical diagnostics has undergone substantial advancements, particularly with the rise of computer-aided diagnostic (CAD) systems. These deep

Manuscript received December 17, 2024; revised February 3, 2025. Date of publication March 25, 2025. Date of current version March 25, 2025. The associate editor prof. Maja Braović has been coordinating the review of this manuscript and approved it for publication.

Alaa M. Youssef (Corresponding author) is with the Faculty of Computers & Artificial Intelligence, Helwan University, Cairo, Egypt (e-mail: ayoussef2002@hotmail.com).

Wessam H. El-Behaidy is with the Faculty of Computers and Artificial Intelligence, Helwan University, Egypt, and the Faculty of Informatics and Computer Science, British University in Egypt (BUE), El-Sherouk, Egypt (e-mails: w_behaidy@fci.helwan.edu.eg, wessam.elbehaidy@bue.edu.eg).

Aliaa A. A.Youssif is with the College of Computing and Information Technology Arab Academy for Science, Technology & Maritime Transport (AASTMT), Cairo, Egypt (e-mail: aliaay@aast.edu).

Digital Object Identifier (DOI): 10.24138/jcomss-2024-0113

learning-driven systems play a vital role in supporting practitioners in the detection and differentiation of various abnormalities [8]. The effectiveness of a CAD system depends on the precise selection of features and learning capabilities of the classifier, making the identification of relevant features crucial for training an accurate classification model.

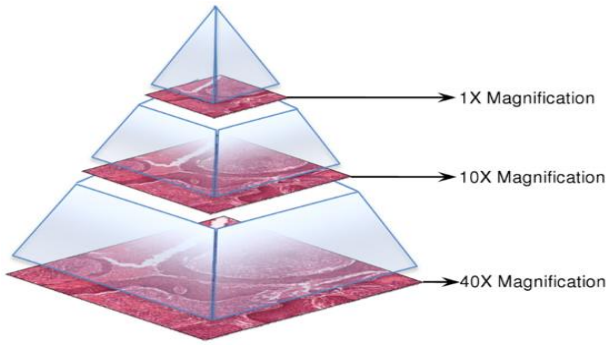

Fig. 1. Typical WSI image with multiscale pyramid resolution [5].

Breast cancer, the second most prevalent lethal disease, poses a substantial challenge in medical imaging, particularly in the classification of mitotic nuclei, which are key indicators of tumor malignancy and cancer progression. To address this challenge, several CAD systems based on machine learning have been devised to automate mitosis classification [9]. However, automating mitosis classification is daunting because of the high resemblance between mitotic and non-mitotic examples, coupled with variations in the appearance of mitotic nuclei [10]. Many existing classification techniques struggle to discern the intrinsic features that differentiate these two classes.

In the broader scope of breast cancer diagnostics, the challenge extends to categorizing cancer levels [11, 12]:

- Benign tumors are noncancerous growths that typically do not invade the surrounding tissues or spread to other parts of the body.
- In situ carcinoma where cancer cells are confined to the site of origin without invading neighboring tissues
- Invasive carcinoma where malignant cells penetrate surrounding tissues and possibly spread to distant parts of the body through the bloodstream or lymphatic system.

Discriminating these diverse levels of cancer requires a nuanced understanding of the distinctive features of each cancer type. Traditional methods often fail to capture these nuances [13].

CNNs have gained prominence because of their ability to learn robust feature representations and effectively classify image data. The feature-learning capabilities of CNNs hinge on both the input representation of images and architectural design of the network [14]. Consequently, enhancing the feature space improves the learning capacity of CNNs for discriminating features, thereby achieving a superior performance in classification tasks. This study delves into the application of CNNs, with a specific focus on improving the feature space in the context of breast cancer segmentation to address the challenges posed by mitosis and cancer-level classification [13, 14].

This study makes several key contributions to the field of medical image segmentation, particularly in the context of breast cancer histopathological images. By enhancing the DeepLabV3+ architecture and integrating advanced techniques, this research aims to improve segmentation accuracy and model stability. The following points summarize the main contributions of this study:

- We propose modifications to the standard DeepLabV3+ architecture by integrating an attention layer between the encoder and decoder to improve model stability and accuracy.
- A spatial attention module was developed to significantly improve the focus of the model on important features, resulting in enhanced segmentation performance, especially in difficult areas of histopathological images.
- Use of the Xception backbone, as we introduce and evaluate the use of the Xception backbone in two modified models, demonstrating its superiority over the traditional ResNet50 backbone in handling complex feature extraction tasks, resulting in higher overall accuracy. Xception significantly minimizes the number of trainable parameters and computations [15].
- Comprehensive performance analysis is conducted by evaluating the proposed models across various metrics, including training and validation loss, learning stability, and accuracy, highlighting the improvements over the standard DeepLabV3+ model.
- Application to breast cancer segmentation involves successfully applying the proposed model to the BCSS dataset, providing a robust solution for segmenting multiple tissue types with varying complexities.

The remainder of this paper is organized as follows. Section II surveys previous research on breast cancer segmentation and detection. Section III discusses the dataset preparation, proposed model, and applied validation measures. Section IV discusses the results. Section V discusses an analysis of the model's performance on the dataset, and Section VI concludes the paper.

II. Related Work

Breast cancer segmentation using histopathological images has garnered significant attention because of its crucial role in accurate diagnosis and treatment planning. Histopathology involves examining tissues under a microscope to obtain detailed information about their cellular morphology and architecture, which is essential for identifying cancerous regions. This process often requires a representation that quantitatively characterizes tumor cells or tissues, relying on the extraction of relevant features to accurately assess tissue and organ function.

Histological images require a quantitative representation of tumor cell or tissue characteristics [16, 17]. The accurate assessment of tissue and organ function depends on capturing the relevant features that define cellular and tissue structures. Feature extraction, which reduces an image to a compressed feature vector, is essential for distinguishing and automatically classifying tissues as normal or malignant as well as for grading them appropriately [17].

Feature extraction can involve handcrafted features, such as identifying critical aspects, including shape, color, and texture, in breast cancer histopathological images. The most significant shape (or morphometric) features of cells include the nuclear

area, convex area, and outline. In cancer cells, morphometric features, such as darker nucleoli, reduced cytoplasm, disorganized chromosomes, abnormal cell growth, and variations in shape and size, are more pronounced. The use of the mean values of these features for the cell and nuclear outlines has been shown to improve the grading process, as demonstrated in [18, 19, 20].

A. Evolution of DeepLab Architectures for Multiscale Feature Extraction

Capturing multiscale features is crucial for the accurate segmentation of histopathological images because of the varying sizes and shapes of cellular structures. Through their evolutionary path, DeepLab architectures have harnessed the formidable capabilities of atrous convolutional layers, thereby playing a pivotal role in their advancement. These layers provide explicit control over the feature resolution within CNNs, effectively enlarging the field of view of the filters without increasing the number of bloating parameters or computational complexity. Their primary objective was to expand the receptive field over the input feature maps while preserving the spatial dimensions and avoiding an explosion in the network parameters [21].

DeepLabV2 [22] introduces multiple parallel atrous convolutional layers with varying dilution rates, collectively forming an atrous spatial pyramid pooling (ASPP) model. The ASPP model facilitates object segmentation across multiple scales and enhances semantic understanding.

DeepLabV3 [23] builds upon this foundation by incorporating image-level features, a 1×1 convolution, and three 3×3 atrous convolutions with distinct dilution rates.

DeepLabV3+ [24] integrates an encoder-decoder structure with the ASPP model. Depth-wise separable convolutions were employed in both the ASPP and the decoder modules. The encoder systematically reduces the spatial dimensions through repeated convolution and pooling layers, whereas the decoder progressively restores spatial information through deconvolution and upsampling layers. To further increase segmentation precision, skip connections were introduced between the encoder and decoder modules.

This trajectory of architectural enhancements underscores the continuous refinement of DeepLab models, leading to more sophisticated and accurate semantic segmentation results [13].

B. Modifications and Applications of the DeepLab Architecture across Diverse Domains

In various studies, the DeepLab architecture has been modified and applied to diverse domains, significantly improving the segmentation accuracy and efficiency. One study incorporated DeepLabV3 into a pipeline for segmenting various histopathological structures, enhancing the model's ability to adapt to different tissue types and improving the overall segmentation quality [25]. Another study modified DeepLabV3+ to include an uncertainty-driven pooling layer, which enhanced the segmentation accuracy and reliability by better handling ambiguous regions in histological images [26].

DeepLabV3+ was employed for liver tumor segmentation from computed tomography (CT) images, utilizing deep convolutional neural networks (DCNNs) combined with probabilistic graphical models (DenseCRFs) to capture multiscale contextual information and refine the target boundaries [27]. In [28], DeepLabV3+ provided superior results to traditional methods such as VGG16, ResNet18, SqueezeNet, and MobileNetV2 for histopathological image segmentation of breast cancer.

For semantic segmentation of buildings in high-resolution remote sensing images, DeepLabV3, which leverages the ability of the model to handle high-resolution inputs and effectively capture multiscale features, was utilized in [29]. In [30], DeepLabV3+ was applied to microscopic breast cancer segmentation, demonstrating its effectiveness in accurately segmenting and classifying cancerous regions in histopathological images.

In [31] DeepLabV3+ was used to segment smooth muscle fibers in hematoxylin and eosin (H&E)-stained images, achieving high accuracy and robustness. In [32], DeepLabV3+ segmented cancerous tissues, benefiting from its encoder-decoder structure, which facilitated detailed and accurate segmentation.

In [33], DeepLabV3+ with an Xception-65 [34] backbone was used to recognize pathological T stages and tumor invasion in rectal cancer, leveraging atrous spatial pyramid pooling for multiscale feature extraction.

In [35], DeepLabV3+ was incorporated into a framework for segmenting cracks on concrete surfaces, highlighting its effectiveness compared to other architectures, such as U-Net and fully convolutional networks (FCN). In [36], DeepLabV3+ outperformed UNet++ [37] in accuracy for oral epithelial dysplasia (OED) segmentation.

Another study by Han et al. [7] utilized DeepLabV3+ to segment breast cancer via the " Breast Cancer Semantic Segmentation" (BCSS) dataset [38]. They introduced a modified ASPP module with varying atrous rates and incorporated a hybrid loss function that combined cross-entropy and Dice losses. These enhancements facilitated better segmentation results, particularly for small and irregularly shaped regions, with an accuracy of 84.83%.

C. Advanced Architectures for Breast Cancer Tissue Segmentation on the BCSS Dataset

Srijay et al. [39] introduced a sophisticated breast cancer tissue-segmentation architecture. The model integrates spatially adaptive normalization layers within a ResNet-based framework to increase the segmentation accuracy of histopathological images. By preserving fine details through these normalization layers, the model addresses the common issue of detail loss in deep learning models during segmentation tasks. The architecture of the model processed image tiles and produced refined segmentation maps, achieving an impressive overall accuracy of 77%.

Huang et al. [40] introduced a two-stage weakly supervised semantic segmentation model emphasizing the relationships between different tissue types in histopathological images. The model employs a graph-parsing attention mechanism to refine the pseudo-labels, which are subsequently used to guide the segmentation process. The authors also introduced a cyclic pseudo-mask strategy and multi-resolution supervision, which further enhanced the segmentation accuracy of the model. This approach allows for a more precise segmentation by leveraging weak supervision to achieve a superior accuracy of 85.94%.

Mauricio et al. [41] introduced an advanced iteration of the traditional U-Net architecture tailored for multiclass breast cancer tissue segmentation. The model features densely connected blocks and residual connections, which enhance the extraction and representation of features from the histopathological images. These architectural improvements enable the model to handle the complexity and diversity of breast cancer tissues effectively, resulting in more accurate segmentation. The model achieved an impressive accuracy of 81%, demonstrating its effectiveness in precise tissue segmentation.

In conclusion, DeepLabV3+ has shown promising results in various domains, including biomedical imaging, histopathology, magnetic resonance imaging (MRI) analysis, and remote sensing, and can be effectively utilized in histopathological image segmentation of breast cancer.

III. Materials and Methods

This section discusses dataset preparation, evaluation metrics, and design of the proposed system.

A. Dataset Preparation

For the dataset preparation phase, the models were trained using the BCSS dataset [38]. This dataset comprises 155 H&E images, each accompanied by 20000 semantic segmentation annotations for tissue regions across 22 classes. To create a focused subset for training, we extracted 57000 sub-images representing cancer regions of interest (ROIs) from the original 155 images. Data augmentation techniques were applied to the sub-images to enhance the model generalization. Additionally, subimages containing classes representing less than 1% of the dataset were removed to avoid class imbalance and reduce noise during the training stage. These subimages, each with a size of 256×256 pixels, were sampled with a 50% overlap for the cancer portions, as shown in Fig. 2. The red and green boxes represent the first and second subimages, respectively, with an overlap of 128 pixels. The resulting dataset was partitioned randomly into 42800 for training, 10700 for validation, and 3500 for testing.

The dataset distribution revealed the classes with the most balanced combination, with the tumor and stroma being the dominant classes as illustrated in Fig. 3 tumors accounted for 42.64%, stroma accounted for 37.78%, lymphocytic infiltration accounted for 10.49%, necrosis or debris accounted for 4.49%, and other types accounted for 4.60%. This imbalance is largely due to the higher prevalence of tumor and stroma images in the dataset. Consequently, even with data augmentation, the ratio of tumor and stroma classes remains high, highlighting the inherent challenge of achieving a perfectly balanced dataset for breast cancer histopathology.

B. Evaluation Metric

The evaluation metrics for the segmentation performance of the proposed detector utilized four widely recognized metrics. These key metrics are based on the counts of true positives (TPs), false positives (FPs), false negatives (FNs), and true negatives (TNs) across all images in the dataset. True positive (TP) instances were identified as the intersections between the segmented cell membrane and its ground truth, whereas those not meeting this criterion were categorized as false positives (FPs). False negatives (FNs) are determined by the missed portions of the ground truth, whereas true negatives (TNs) encompass the image regions beyond the union of the segmentation and ground truth.

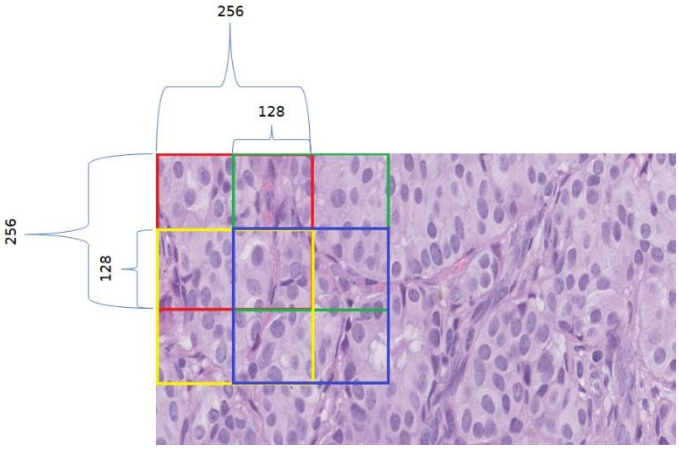

Fig. 2. Process of the moving window with 50% overlap

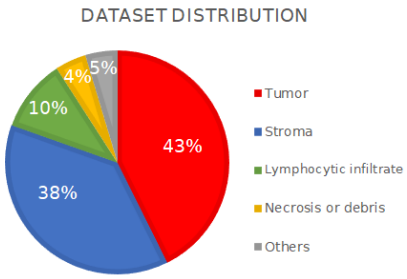

Fig. 3. Dataset class distribution

B.1 Network Accuracy Metric

It is a performance measure that quantifies the overall accuracy of the predictions made by a CNN. It is a comprehensive metric that considers both true-positive and true-negative predictions relative to the total number of instances in the dataset [42]. The network accuracy can be determined using

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

In the context of CNNs, accuracy is a fundamental metric that assesses the ability of a model to correctly classify instances across all the classes. This approach provides a general overview of the model's effectiveness but may not be suitable for highly imbalanced datasets. In such cases, additional metrics such as precision, recall, and F1 score are often considered to provide a more nuanced evaluation of the model's performance.

B.2 F1 Score Metric

This is a measure of a model's accuracy that balances both precision and recall [43]. It provides a single numerical value that considers both false positives and false negatives, thereby offering a comprehensive evaluation of the performance of a model. The F1 score can be determined using

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

where

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

and

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

B.3 Mean Intersection over Union (MIoU)

Mean intersection over union (MIoU) is a common evaluation metric used for segmentation tasks in computer vision [44]. It measures the overlap between the predicted segmentation and the ground truth, providing an overall performance score. The MIoU can be determined using

$$\text{MIoU} = \frac{1}{C}\sum_{i=1}^{C}\frac{P_i\cap G_i}{P_i\cup G_i} \quad (3)$$

where P_i is the predicted set of pixels for class i, G_i is the ground truth set of pixels for class i and C is the total number of classes. The intersection represents the common pixels between the prediction and ground truth, whereas the union represents the total number of pixels present in either the prediction or the ground truth. MIoU is the average IoU value for all classes.

B.4 Frequency weighted intersection over union (FwIoU)

The frequency-weighted intersection over union (FwIoU) is a variant of the IoU metric, which considers the frequency of each class in the dataset [44]. This metric gives more importance to classes that appear more frequently in datasets. The FwIoU can be determined using

$$\text{FwIoU} = \sum_{i=1}^{C}\frac{n_i}{N}\cdot\frac{P_i\cap G_i}{P_i\cup G_i} \quad (4)$$

where n_i is the number of pixels belonging to class i and where N is the total number of pixels across all classes. This metric weighs the IoU of each class by its proportion in the dataset, thereby providing a balanced evaluation that considers the class imbalance.

C. The Proposed Model

We introduced CancerSeg-XA, a semantic segmentation model specifically tailored to breast cancer histopathology images. We used the advanced DeepLabV3+ architecture and three modified versions to enhance segmentation performance. The first version, Model 1, integrates an attention layer between the encoder and decoder of the DeepLabV3+. The second version, Model 2, employs the Xception backbone up to "block5" instead of the original ResNet50. In the final version, CancerSeg-XA, we combined the Xception backbone with an attention layer between the encoder and the decoder. The model architectures are shown in

Fig. 4. Semantic segmentation of the histopathological images of breast cancer is essential for accurate diagnosis and treatment planning.

Our models aimed to precisely segment various tissue structures and cell types associated with breast cancer by leveraging multiscale contextual information, spatial details, and attention mechanisms. Building on the robust framework of DeepLabV3+, our approach addresses the unique challenges of breast cancer segmentation to achieve a superior performance. In the subsequent sections, we provide a detailed description of the model architectures, highlighting key components such as the encoder, ASPP, decoder, skip connections, and the integrated attention layer.

C.1 Encoder

The encoder layer in the models receives an input subimage with dimensions of 256×256×3 for feature extraction. In Model 1, we utilized the encoder of the original DeepLabV3+ architecture, which employed the ResNet50 [45] backbone. This architecture uses four residual unit blocks (Res1–Res4) to extract low-level features, outputting 64 channels used as inputs to the next layer.

For Model 2 and CancerSeg-XA, we replace the ResNet50 backbone with the Xception architecture, which extends up to "block5." This modification leveraged the depth-wise separable convolutions of Xception to capture more intricate spatial details. Using different backbones, we aimed to explore the impact of feature extraction capabilities on segmentation performance.

These diverse encoder configurations allowed us to investigate the effectiveness of different architectures in extracting low-level features that are crucial for precise segmentation in breast cancer histopathological images.

C.2 Atrous spatial pyramid pooling (ASPP)

Four parallel atrous separable convolutions (ASC) with diverse dilation rates are employed on the output from the encoder backbone, enabling the analysis of extracted features across various scales. Atrous separable convolution (ASC) comprises depthwise convolution with atrous convolutions, followed by pointwise convolution. This architecture facilitates the examination of feature details for multiple receptive field sizes. In addition,, an average pooling layer is incorporated to further refine the feature representations. Consequently, the ASPP module produces an output of 1280 channels, providing comprehensive feature representations suitable for the subsequent process.

DeepLabV3+ uses the ResNet50 backbone to maintain robust feature-extraction capabilities. However, Xception architecture, with its depth-wise separable convolutions, significantly enhances the model's ability to capture intricate spatial details while maintaining efficiency. This modification results in improved feature extraction, allowing the ASPP module to generate richer and more detailed feature maps. This combination ensures that our model can effectively handle the complex and varied structures found in histopathological images of breast cancer, thereby delivering robust feature representations for precise segmentation.

The transition from ResNet50 to Xception not only retains the high performance of the original architecture but also provides additional benefits in terms of computational efficiency and feature richness, as demonstrated in various studies on image segmentation tasks. This adaptation is crucial for achieving a superior segmentation performance in the challenging domain of breast cancer histopathological images.

C.3 Attention Layer

An attention layer block was introduced between the encoder and decoder layers to augment channel priority. This layer significantly enhances the network's ability to focus on the most relevant features by employing a spatial attention module (SAM).

SAM takes an input tensor with dimensions of 16×16×256 and performs a convolution operation that outputs a 16×16×1 tensor.

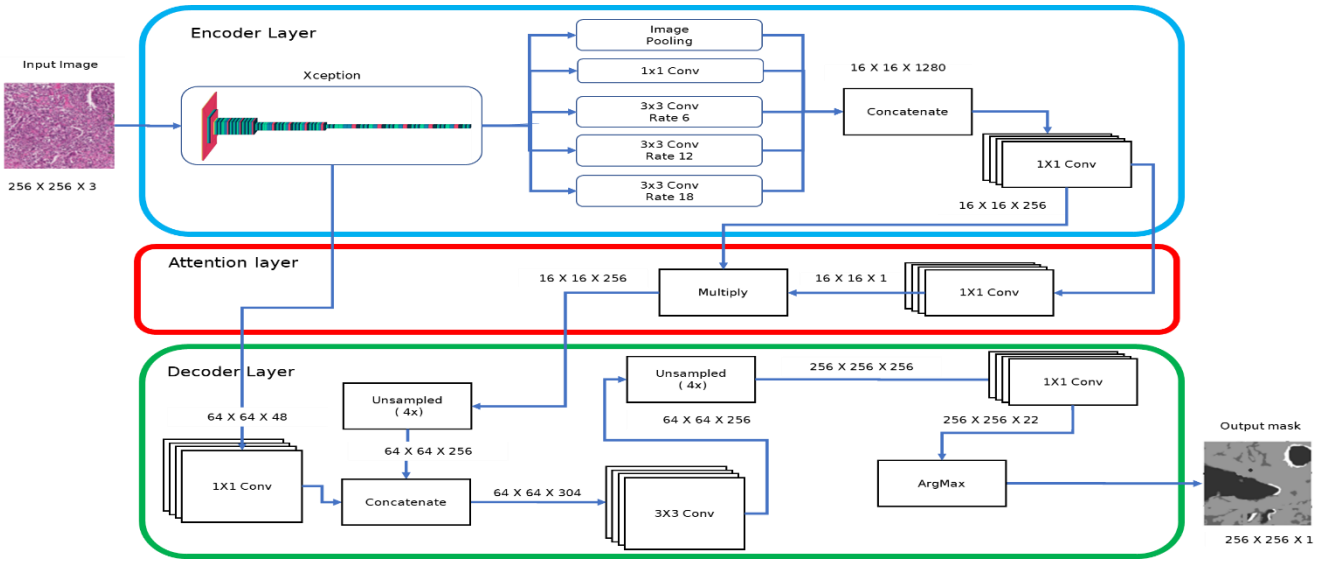

Fig. 4. The proposed system modifications based on DeepLabV3+

This spatial attention map highlights the crucial spatial features by applying a sigmoid activation function. The resultant attention map was then multiplied element-wise by the original input tensor, resulting in an output tensor with dimensions of 16×16×256. This process helps emphasize the important spatial regions of the input tensor while suppressing the less relevant regions.

By incorporating these attention mechanisms, the model can dynamically adjust the importance of features, thereby enhancing the segmentation accuracy and robustness, particularly in the context of breast cancer histopathological images.

C.4 Decoder

In the decoder stage, two inputs are processed: the first input is feature maps of size 64×64×48 and the second input is from the Res2 block, which has dimensions of 16×16×256. The second input was upsampled to 64×64×256 pixels. These two inputs were concatenated and subjected to a 3×3 convolution to reduce the output to 256 channels. The output was then upsampled by a factor of four, and a final 1×1 convolution was applied to produce a pixel-wise segmentation mask that represents the segmentation prediction for each pixel in the input image.

The output of the upsampling operation is restored to the original input image size of 256×256×22, representing the weights for the 22 classes. The ArgMax function was then applied to obtain a final mask of size 256×256, representing the segmentation prediction for each pixel in the input image.

IV. Results Analysis

This section specifies the hyperparameters used during training, the performance of DeepLabV3+ and the three modified models, and finally the impact of architectural modifications on model performance.

A. Training Setup

CancerSeg-XA was implemented in Python 3.7 via the Keras library as an interface for TensorFlow. The training was performed in the Kaggle environment, which provides a P100 GPU with 16 GB of RAM and powerful cloud-computing resources. The model used the Adam optimizer and sparse categorical cross-entropy loss function.

A reduced learning rate function was used to adjust the learning rate, with a reduction factor of 0.2 and a patience parameter set to five epochs. The CNN models trained by tuning hyperparameters are shown in Table I.

TABLE I
MODEL TRAINING HYPERPARAMETER VALUES

Hyperparameter	DeepLabV3+	Model 1	Model 2	CancerSeg-XA
Initial learn rate	1e-3			
End learn rate	1e-4	1e-3	1e-5	
No. of epochs	25		45	
Batch size	24			
Image size	256 x 256			
Trainable params	10611830	10613111	8037198	8038479
Nontrainable params	32224	32224	13456	13456

B. Training Progress

In this section, we evaluate the performance of DeepLabV3+ and the three modified models along with their training performance. Fig. 5 shows the training validation loss, generalization ability, and overfitting resistance of the models. For DeepLabV3+, the first five epochs showed that the model struggled with a learning rate (LR) of 1e-3, but the stability improved once the LR was decreased by a factor of 0.2 to 2e-4. In Model 1, the attention layer overcomes early instability, resulting in greater learning stability and improvement by the end of 25 epochs. Model 2 and CancerSeg-XA, which utilized the Xception backbone, displayed a stable start with loss values of less than 1. Model 2 required a reduction in LR to 2e-4 at epoch 12 and further to 1e-5 at epoch 31. CancerSeg-XA, enhanced by the attention layer, demonstrated improved learning stability, requiring an LR decrease only at epochs 27 to 2e-4.

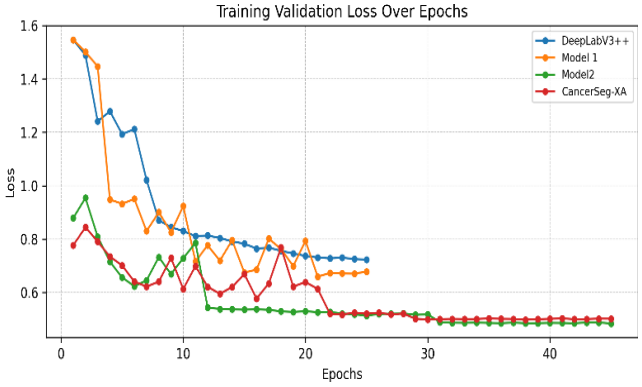

Fig. 5. Result analysis for model training validation loss

The accuracy of the training validation is shown in Fig. 6. The training validation accuracy results showed that Model 1 achieved higher accuracy than DeepLabV3+. This improvement is due to the added attention mechanism. However, Model 2 and CancerSeg-XA, which incorporated the Xception backbone, outperformed DeepLabV3+ and Model 1 in terms of accuracy. The superior feature extraction capabilities of the Xception backbone contributed significantly to this performance increase. While the attention layer in Model 1 and Model 4 slightly improved the final result, its primary benefit was enhanced training stability.

C. Results

To the best of our knowledge, only a limited number of studies have utilized DeepLabV3+ for the segmentation of histopathological images of cancer. Therefore, to validate CancerSeg-XA and assess its performance, we selected DeepLabV3+ as a baseline model benchmark for CancerSeg-XA. After a detailed examination and performance evaluation, CancerSeg-XA demonstrated substantial improvements over the baseline model accuracy by 5.86%, recall by 10.03%, F1-score by 10.48%, MIou by 8.48%, and FwIou by 10.88%, as shown in Table II.

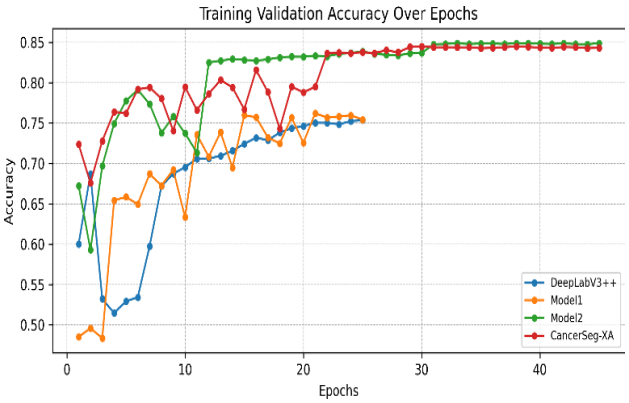

Fig. 6. Result analysis for model training validation accuracy

TABLE II
PROPOSED MODEL RESULTS

Model	Accuracy	Recall	F1-score	MIou	FwIou
Han et al. [7]	0.84832	-	-	0.6892	0.7374
Srijay et al. [39]	0.77	-	-	-	-
Huang et al. [40]	0.8594	-	-	0.6989	-
Mauricio et al. [41]	0.81	-	-	-	-
DeepLabV3+	0.8570	0.7926	0.7834	0.5166	0.7312
Model 1 (Resnet50 backbone + Attention layer)	0.8694	0.801	0.7928	0.51	0.7374
Model 2 (Xception backbone)	0.9147	0.8901	0.8867	0.593	0.8361
CancerSeg-XA (Xception backbone + Attention layer)	0.9157	0.8929	0.8883	0.6014	0.84

The CancerSeg-XA was evaluated against popular methods, including those by Han et al. [7] and Srijay et al. [39], Huang et al. [40], and Mauricio et al. [41]. CancerSeg-XA showed notable improvements in most metrics. Specifically, it achieved a 14.75% greater accuracy than Srijay et al. [39], a 10.57%

greater accuracy than Mauricio et al. [41], a 6.54% greater accuracy than Han et al. [7], a 5.63% greater accuracy than did Huang et al. [40], and a 10.26% greater FwIoU than did Han et al. [7].

Huang et al. [40] demonstrated a modest improvement over the CancerSeg-XA in terms of the MIoU metric, with a 9.75% higher score. This superior performance is likely attributed to the use of a multiresolution supervision strategy, which effectively narrows the gap between the image-level and pixel-level labels. Notably, their model achieved an MIoU of 0.8335 before implementing this strategy. The multiresolution supervision approach appears to be particularly advantageous for preserving the intersection over the union across multiple classes, especially in scenarios involving more complex and overlapping structures.

D. Impact of Architectural Modifications on Model Performance

The performance evaluation of the proposed changes is summarized in Table III, highlighting the impact of replacing the ResNet50 backbone with the Xception backbone (up to Block5)

and the integration of an attention layer. As shown, the transition to the Xception backbone led to a significant reduction in model complexity by reducing the number of layers from 147 (DeepLabV3+) to 52 (CancerSeg-XA). This architectural change also led to a significant reduction in model size by 24.2%, from 40.48 MB to 30.66 MB.

The integration of the attention layer, while adding only two layers and 1281 trainable parameters (less than 0.01% of the total parameters), demonstrated an excellent balance between minimal overhead and performance gains. These modifications collectively improve the segmentation accuracy and model stability, reinforcing the effectiveness of the proposed design while maintaining computational efficiency.

TABLE III
COMPARISON OF THE MODEL'S PERFORMANCE AND ARCHITECTURAL CHARACTERISTICS

Model	Time cost /Step	No. of layers	No. of trainable param	Size
DeepLabV3+	391 ms	147	10611830	40.48 MB
Model 1 (Resnet50 backbone + Attention layer)	378 ms	149	10613111	40.49 MB
Model 2 (Xception backbone)	367 ms	50	8037198	30.66 MB
CancerSeg-XA (Xception backbone + Attention layer)	369 ms	52	8038479	30.66 MB

These optimizations enhance the computational efficiency and scalability of the model, making it more suitable for real-world deployment scenarios, including environments with limited resources.

V. DISCUSSION

In this section, we discuss the performance of DeepLabV3+ and our modified version, particularly when analyzing the models' performance on images belonging to a single class, as shown in Fig. 7, Models 2 and CancerSeg-XA demonstrated equal or superior accuracy compared with DeepLabV3+ and Model 1. This performance enhancement can be attributed to the architectural improvements and efficient feature extraction capabilities provided by the Xception backbone in Model 2 and CancerSeg-XA.

For images belonging to multiple classes, the Xception-backed models achieved significantly better accuracy, as illustrated in Fig 8. This improvement is likely due to the superior ability of Xception to extract local features, which is critical for distinguishing between different tissue types and structures in complex histopathologic images. The improved feature extraction capabilities of Xception enable better segmentation and classification, leading to more accurate predictions in multiclass scenarios. DeepLabV3+ achieves the best accuracy when image classes are restricted to tumors and stroma, and has lower accuracy on images containing other classes. This highlights the importance of the attention layer in Models 1 and CancerSeg-XA, where the segmentation accuracy improves significantly compared with DeepLabV3+ and Model 2, respectively.

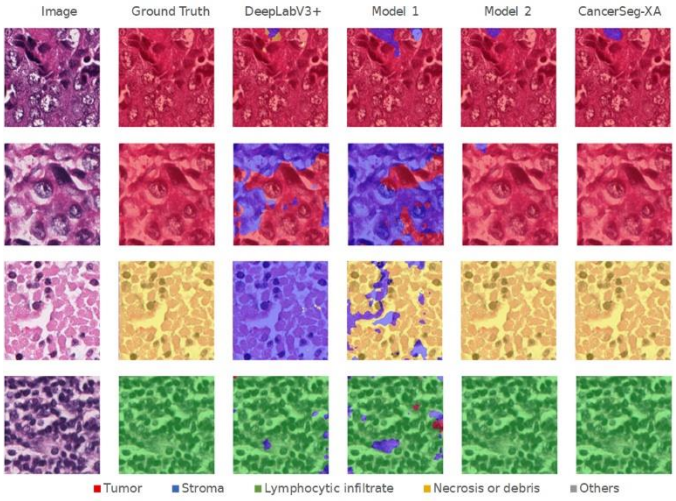

Fig. 7. Model performance for a single class

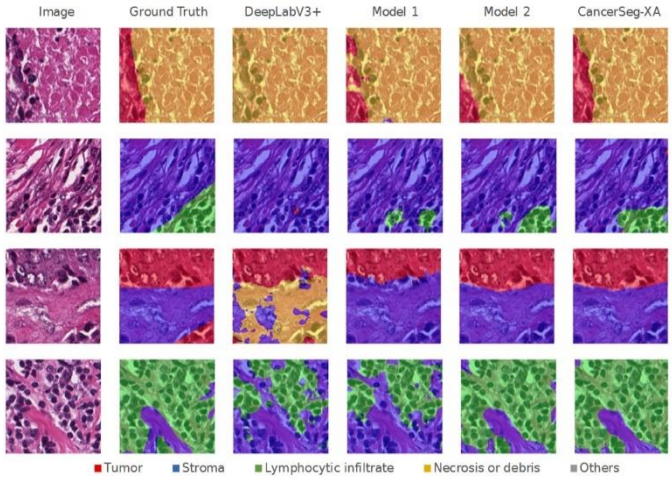

Fig. 8. Performance for multiclass images

Training and segmentation performance are significantly influenced by dataset imbalance, particularly when examining class distributions within the BCSS dataset. As "tumor" and "stroma" are the dominant classes, their segmentation performance remained largely unaffected by class imbalance, as the model was able to learn sufficient discriminative features. However, minority classes such as "lymphocytic infiltration" and "necrosis or debris" posed greater challenges due to their limited representation. Despite this, CancerSeg-XA successfully identified these underrepresented classes with acceptable accuracy, demonstrating the model's robustness in learning meaningful features even from sparse data as shown in Fig. 7 and Fig 8. A more critical issue arises when an image contains a majority of the "Other" class, which typically accounts for only 5% of the overall dataset but can dominate certain images. Since this class lacks distinct structural patterns, it significantly affects segmentation performance, making it difficult for the model to differentiate clinically relevant regions from ambiguous background information.

TABLE IV
COMPARISON OF SEGMENTATION APPROACHES ON THE BCSS DATASET

Model	Approach	Accuracy	MIou
Sun et al. [46]	Semi-supervised learning	0.8398	0.71
MSF-WSI [47]	Semi-supervised Learning	0.9136	0.647
DETisSeg [48]	Vision transformers	0.7986	0.6491
Bi-VLGM [49]	Vision transformers	0.9359	0.703

To further explore the effectiveness of CancerSeg-XA, which leverages fine-tuning of pre-trained models, we compare its results with other state-of-the-art approaches, as shown in Table IV. The comparison includes semi-supervised learning methods such as Sun et al. [46] and MSF-WSI [47], and vision transformer-based architectures like DETisSeg [48] and Bi-VLGM [49]. While vision transformers may achieve slightly higher accuracy, they come with a significant trade-off in terms of computational complexity and training time. DETisSeg [48] requires 320000 iterations, while Bi-VLGM [49] takes 40000 iterations, making them computationally expensive. Similarly, semi-supervised learning methods demand 300 epochs for training, adding to the overall resource burden. In contrast, CancerSeg-XA achieves competitive performance with only 45 epochs, demonstrating a more balanced trade-off between accuracy and efficiency. This highlights the advantage of fine-tuning pre-trained models, allowing for high segmentation performance while maintaining significantly lower computational costs, making it more practical for real-world medical applications.

VI. CONCLUSION

Breast cancer remains a major global health problem that requires reliable computational tools for its diagnosis and treatment. This study demonstrates the potential of the DeepLabV3+ model and its modified versions for semantic segmentation of histopathologic images of breast cancer. Our results show that the inclusion of an attention layer increases the training stability, whereas replacing the ResNet50 backbone with the Xception backbone significantly improves segmentation accuracy. Among the models tested, CancerSeg-XA achieved the highest accuracy of 91.57%, an improvement of 5.86% over the baseline DeepLabV3+ model. In addition, CancerSeg-XA has a lower complexity, size, and number of layers than DeepLabV3+, as it utilizes the Xception backbone to focus on local features that are critical for distinguishing between different tissue types and structures in complex histopathology images. These results highlight the importance of advanced architectures and attention mechanisms to improve segmentation performance and pave the way for more accurate, efficient, and robust models for clinical applications.

Future work will focus on building on the promising results of the CancerSeg-XA model, which combines high performance with a simplified complexity. This includes exploring more advanced attention mechanisms to further improve segmentation accuracy and robustness. In addition, the adaptability of the model can be tested on datasets such as PanNuke to extend its application to different cancer types and to demonstrate its potential for wider use in medical image analysis.

REFERENCES

[1] "WHO - Cancer," [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cancer. [Accessed 14 02 2025].

[2] "Global cancer incidence in women," American Institute for Cancer Research, 2018. [Online]. Available: https://www.wcrf.org/dietandcancer/worldwide-cancer-data/. [Accessed 14 02 2025].

[3] A. Ponraj and A. Canessane, "Radial Basis Function Networks and Contrast-Limited Adaptive Histogram Equalization Filter Based Early-Stage Breast Cancer Detection Techniques," *Journal of Computer Science*, pp. 760-774, 2023.

[4] S. R. Duenweg, S. A. Bobholz, A. K. Lowman, M. A. Stebbins, A. Winiarz, B. Nath, F. Kyereme, K. A. Iczkowski and P. S. LaViolette, "Whole slide imaging (WSI) scanner differences influence optical and computed properties of digitized prostate cancer histology," *Journal of Pathology Informatics*, vol. 14, 23.

[5] Y. Wang, K. E. Williamson, P. J. Kelly, J. A. James and P. W. Hamilton, "SurfaceSlide: A Multitouch Digital Pathology Platform," *PloS one*, vol. 7, no. 1, 2012.

[6] D. Neofytos, A. Ognjen and C. P. D., "Deep Learning for Whole Slide Image Analysis: An Overview," *Frontiers in Medicine* , vol. 6, p. 264, 2019.

[7] C. Han, J. Lin, J. Mai, Y. Wang, Q. Zhang, B. Zhao, X. Chen, X. Pan, Z. Shi, Z. Xu, S. Yao, L. Yan, H. Lin, X. Huang, C. Liang, G. Han and Z. Liu, "Multi-layer pseudo-supervision for histopathology tissue semantic segmentation using patch-level classification labels," *Medical Image Analysis*, vol. 80, 2022.

[8] K. Doi, "Computer-aided diagnosis in medical imaging: historical review, current status and future potential," *Comput Med Imaging Graph*, vol. 31, pp. 198-211, 2007.

[9] Y. Jiang, R. M. Nishikawa, R. A. Schmidt, C. E. Metz, M. L. Giger and K. Doi, "Improving breast cancer diagnosis with computer-aided diagnosis," *Academic Radiology*, vol. 6, no. 1, pp. 22-33, 1999.

[10] C. Li, X. Wang, W. Liu, L. J. Latecki, B. Wang and J. Huang, "Weakly supervised mitosis detection in breast histopathology images using concentric loss," *Medical Image Analysis*, vol. 53, pp. 165-178, 2019.

[11] X. Li, M. Radulovic, K. Kanjer and K. N. Plataniotis, "Discriminative Pattern Mining for Breast Cancer Histopathology Image Classification via Fully Convolutional Autoencoder," *IEEE Access*, vol. 7, pp. 36433-36445, 2019.

[12] S. Vesal, N. Ravikumar, A. Davari, S. Ellmann and A. Maier, "Classification of breast cancer histology images using transfer learning," *Springer, Cham*, p. 812–819, 2018.

[13] R. Krithiga and P. Geetha, "Breast Cancer Detection, Segmentation and Classification on Histopathology Images Analysis: A Systematic Review," *Archives of Computational Methods in Engineering*, vol. 28, p. 2607–2619, 2020.

[14] A. Khan, A. Sohail, U. Zahoora and A. S. Qureshi, "A Survey of the Recent Architectures of Deep Convolutional Neural Networks," *Artificial Intelligence Review* , vol. 53, p. 5455–5516, 2020.

[15] N. A. El-Mawla, M. A. Berbar, N. A. El-Fishawy and M. A. El-Rashidy, "A novel deep learning approach (Bi-xBcNet-96) considering green AI," *Neural Computing and Applications*, vol. 36, pp. 12701-12723, 2024.

[16] J. de Matos, A. de Souza Britto Jr., L. E. S. Oliveira and A. L. Koerich, "Histopathologic Image Processing: A Review," *Computer Vision and Pattern Recognition*, 2019.

[17] A. Das, M. S. Nair and S. D. Peter, "Computer-Aided Histopathological Image Analysis Techniques for Automated Nuclear Atypia Scoring of Breast Cancer: a Review," *Journal of Digital Imaging*, vol. 33, pp. 1091-1121, 2020.

[18] M. M. Dundar, S. Badve, G. Bilgin, V. Raykar, R. Jain, O. Sertel, and M. N. Gurcan, "Computerized classification of intraductal breast lesions using histopathological images," *IEEE Trans Biomed Eng* , vol. 58, no. 7, p. 1977–1984, 2011.

[19] D. Kolarević, T. Vujasinović, K. Kanjer, J. Milovanović, N. Todorović-Raković, D. Nikolić-Vukosavljević, and M. Radulovic, "Effects of different preprocessing algorithms on the prognostic value of breast tumor microscopic images," *Microsc*, vol. 240, no. 1, p. 17–26, 2018.

[20] I. Prvulović, I. Kardum-Skelin, D. Sustercić, J. Jakić-Razumović, and S. Manojlović, "Morphometry of tumor cells in different grades and types of breast cancer," *Colleg Antropolog*, vol. 34, no. 1, p. 99–103, 2010.

[21] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy and A. L. Yuille, "Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs," *ICLR*, 2015.

[22] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence* , vol. 40, no. 4, pp. 834 - 848, 2018.

[23] L.-C. Chen, G. Papandreou, F. Schroff and H. Adam, "Rethinking Atrous Convolution for Semantic Image Segmentation," *ArXiv*, 2017.

[24] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff and H. Adam, " Encoder-decoder with Atrous separable convolution for semantic image segmentation," *the European Conference on Computer Vision*, pp. 833-851, 2018.

[25] D. Sitnik and I. Kopriva, "LEFM-Nets: Learnable Explicit Feature Map Deep Networks for Segmentation of Histopathological Images of Frozen Sections," *arXiv preprint arXiv:2204.06955*, 2022.

[26] M. M. Fraz, M. Shaban, S. Graham, S. A. Khurram, and N. M. Rajpoot, "Uncertainty Driven Pooling Network for Microvessel Segmentation in Routine Histology Images," *Computational Pathology and Ophthalmic Medical Image Analysis: First International Workshop, COMPAY 2018*, pp. 156-164, 2018.

[27] Y. Sun and C. Shi, "Liver Tumor Segmentation and Subsequent Risk Prediction Based on Deeplabv3+," *IOP Conference Series: Materials Science and Engineering*, vol. 612, no. 2, 2019.

[28] S. K. Subramanya, R. Li, Y. Wang, H. Miyamoto and F. Cui, "Deep learning for histopathological segmentation of smooth muscle in the urinary bladder," *BMC Medical Informatics and Decision Making*, vol. 23, no. 122, 2023.

[29] H. Li, J. Zhang, J. Wang, Z. Feng, B. L. ORCID, N. Xiong, J. Zhang, Xiaoting, Y. Li and S. Lin, "Extracting Citrus in Southern China (Guangxi Region) Based on the Improved DeepLabV3+ Network," *Remote Sensing*, vol. 15, no. 23, 2023.

[30] J. Amin, M. Sharif, S. L. Fernandes, S.-H. Wang, T. Saba and A. R. Khan, "Breast microscopic cancer segmentation and classification using unique 4-qubit-quantum model," *Microscopy Research and Technique*, vol. 85, no. 5, pp. 1926-1936, 2022.

[31] S. K. Subramanya, "Deep Learning Models to Characterize Smooth Muscle Fibers in Hematoxylin and Eosin Stained Histopathological Images of the Urinary Bladder," in *Rochester Institute of Technology*, 2021.

[32] J. Musulin, D. Štifanić, A. Zulijani, T. Ćabov, A. Dekanić and Z. Car, "An Enhanced Histopathology Analysis: An AI-Based System for Multiclass Grading of Oral Squamous Cell Carcinoma and Segmenting of Epithelial and Stromal Tissue," *Cancers*, vol. 13, no. 8, p. 1784, 2021.

[33] Y. Ju, L. Zheng, P. Zhao, F. Xin, F. Wang, Y. Gao, X. Zhang, D. Wang and Y. Lu, "Artificial intelligence recognition of pathological T stage and tumor invasion in rectal cancer based on large panoramic pathological sections," *Intelligent Medicine*, vol. 2, no. 3, pp. 141-151, 2022.

[34] F. Chollet, "Xception: Deep Learning With Depthwise Separable Convolutions," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1251-1258, 2017.

[35] P. N. Hadinata, D. Simanta, L. Eddy and K. Nagai, "Crack Detection on Concrete Surfaces Using Deep Encoder-Decoder Convolutional Neural Network: A Comparison Study Between U-Net and DeepLabV3+," *Journal of the Civil Engineering Forum*, vol. 7, no. 3, 2021.

[36] Y. Liu, E. Bilodeau, B. Pollack and K. Batmanghelich, "Automated detection of premalignant oral lesions on whole slide images using convolutional neural networks," *Oral Oncology*, vol. 134, 2022.

[37] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh and J. Liang, "U-Net++: a nested U-Net architecture for medical image segmentation," *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support*, vol. 11045, pp. 3-11, 2018.

[38] M. Amgad, H. Elfandy, H. Hussein, L. A. Atteya, M. A. T. Elsebaie, L. S. A. Elnasr, R. A. Sakr, H. S. E. Salem, A. F. Ismail, A. M. Saad, J. Ahmed, M. A. T. Elsebaie, D. Manthey, D. A. Gutman and L. A. D. Cooper, "Structured crowdsourcing enables convolutional segmentation of histology images," *Bioinformatics*, vol. 35, no. 18, pp. 3461-3467, 2019.

[39] S. Deshpande and D. Parkhi, "SPADESegResNet: Harnessing Spatially-Adaptive Normalization for Breast Cancer Semantic Segmentation," *Annual Conference on Medical Image Understanding and Analysis*, pp. 344-356, 2024.

[40] C. Huang, S. Liao and Y. Peng, "A Two-stage Weakly Supervised Semantic Segmentation Model Based on Pathological Tissue Relationships," *Authorea*, 2023.

[41] M. A. Ortega-Ruíz, C. Karabağ, E. Roman-Rangel, and C. C. Reyes-Aldasoro, "DRD-UNet, a UNet-Like Architecture for Multi-Class Breast Cancer Semantic Segmentation," in *IEEE Access*, vol. 12, pp. 40412-40424, 2024.

[42] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861-874, 2006.

[43] C. J. V. Rijsbergen, in *Information Retrieval*, Butterworths, 1979.

[44] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez and J. Garcia-Rodriguez, "A Review on Deep Learning Techniques Applied to Semantic Segmentation," *arXiv preprint arXiv:1704.06857*, 2017.

[45] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, 2016.

[46] K. Sun, Y. Z. X. Yang, X. Chen and W. Jia, "Semi-supervised breast cancer pathology image segmentation based on fine-grained classification guidance," *Medical & Biological Engineering & Computing*, pp. 901-912, 2024.

[47] H. Wang, E. Ahn and J. Kim, "A multi-resolution self-supervised learning framework for semantic segmentation in histopathology," *Pattern Recognition*, vol. 155, 2024.

[48] P. He, A. Qu, S. Xiao and M. Ding, "DETisSeg: A dual-encoder network for tissue semantic segmentation of histopathology image," *Aiping Qu;Shuomin Xiao;Meidan Ding*, vol. 87, 2024.

[49] W. Chen, J. Liu, T. Liu and Y. Yuan, "Bi-VLGM: Bi-Level Class-Severity-Aware Vision-Language Graph," *International Journal of Computer Vision*, 2024.

Alaa Mohamed Youssef is currently pursuing a Ph.D. degree in Computer Science at Helwan University. She holds both a B.S. degree and an M.S. degree in Computer Science. Her research journey has been marked by an exciting focus on object detection in images, a field where she has made significant contributions. As she progresses in her academic career, she is now keen to deepen her expertise in machine learning and explore emerging fields of medical image analysis.

Wessam Hassan El Behadidy earned her honors bachelor's degree and a Master of Science (M.Sc.) degree in Computer Science from Helwan University (Faculty of Computers and Artificial Intelligence), Cairo, Egypt. Her M.Sc. Her work was on employing AI and machine learning techniques to recognize persons from their voices, and her work was published at one powerful conference: the National Radio Science Conference (NRSC). She completed her Ph.D. in 2012 in Computer Science from Helwan University (Faculty of Computers and Artificial Intelligence), Cairo, Egypt. Her goal was to analyze one of the hepatitis C virus (HCV) proteins, especially genotype 4A, which is common in Egypt, and to predict its three-dimensional structure, specifically the active areas on its surface. Four publications were published in international journals and conferences outside their work. She is currently an Associate Professor at the Department of Computer Science as well as the vice dean for Community and Environmental Affairs at the faculty, and Associate Professor, Faculty of Informatics and Computer Science, British University in Egypt (BUE), El-Sherouk, Egypt. Her research

interests include medical image analysis, protein structure prediction, pattern recognition, and computer vision.

Aliaa Abdel-Haleim Abdel-Razik Youssif is currently a professor at the College of Computing and Information Technology, Arab Academy for Science Technology and Maritime Transport (AAST), Cairo, Egypt. She received her BSc and MSc degrees in telecommunications and electronics engineering from the Helwan University. She also completed her Ph.D. at Helwan University. She had a postdoctoral degree at George Washington University, USA in 2005. In addition, she was a visiting professor at many universities, such as Cardiff University in the United Kingdom (2008), International Telematic University in Italy (2012), Masaryk University, Czech Republic (2016), Oviedo University, Spain (2017), Vilnius Gediminas Technical University, Lithuania (2018), and Cardiff Metropolitan University, UK (2020).