

Invasive Ductal Carcinoma (IDC) Nuclei Classification using Mask RCNN

Amany Ibrahim, Hanaa Torkey, and Ayman El-Sayed

Abstract— Breast cancer is the second most prevalent cancer globally and remains one of the leading causes of cancer-related mortality. Over the past few decades, the incidence of breast cancer has increased significantly, highlighting the critical need for early detection to improve survival rates. In response, researchers have been actively developing computer-aided diagnostic systems to assist in rapid and accurate diagnosis. Various datasets have been utilized in these efforts, leveraging the power of Artificial Intelligence (AI) to support radiologists in medical image analysis, ultimately enhancing patient diagnosis and treatment. Among the available diagnostic techniques, histopathology imaging remains the gold standard for detecting breast cancer with high accuracy. In this study, we employed ResNet- based architectures to implement a Mask Region-based Convolutional Neural Network (Mask R-CNN) for the automated detection of nuclei in histopathological breast cancer images. Following detection, the system classifies the cancer type, extracting multi-scale features using a combination of Feature Pyramid Networks (FPN) modules. To further enhance recognition accuracy, we utilized Region of Interest Align (RoIAlign), ensuring precise feature extraction. Experimental results demonstrate that our proposed approach not only delivers superior visual interpretability but also outperforms existing models in key performance metrics, achieving 97.7% accuracy, 97% recall, and a 96.7% F1 score.

Keywords—FPN , MaskRCNN , Nuclei Segmentation, ResNet.

I. INTRODUCTION

Breast cancer is the most common cancer among women. After lung cancer, breast cancer is the second deadliest cancer according to World

Health Organization (WHO)[1]. The International Agency for Research on Cancer (IARC) provides key insights into the global cancer burden. Here are some highlights from their 2022 estimates: New Cases: Around 20 million new cases of cancer were reported. Deaths: Approximately 9.7 million people died from cancer. Survivors: About 53.5 million people

were alive five years after their diagnosis. Mortality Rates: One in nine men and one in twelve women die from cancer. The IARC's Global Cancer Observatory also offers insights into specific cancer types for 2022. They record 12.4% new cases for Lung Cancer which is approximately 2.5 .11.6% new female cases for Breast Cancer which is about 2.3 million .9.6% new cases for Colorectal Cancer which is about 1.9 million. 7.3% new cases for Prostate Cancer that is about 1.5 million. 4.9% new cases for Stomach Cancer which is 970,000 as shown in figure (1) [1]. Significant progress has been made in cancer treatment. Invasive ductal carcinoma (IDC) may not cause symptoms initially. Often, a mammogram detects abnormalities, prompting further tests. Typically, the first sign of IDC is a lump or mass in the breast, which can be felt by the patient or doctor.

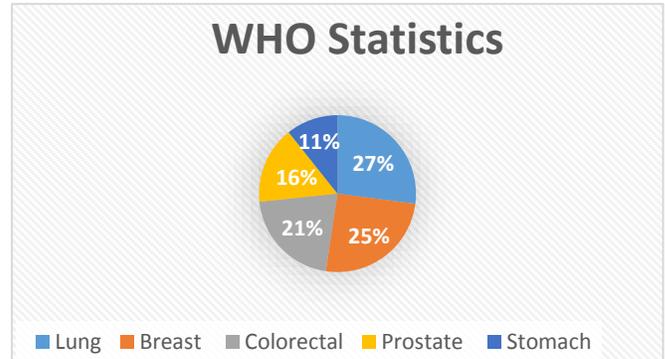


Fig. 1 Who Statistics[1]

The American Cancer Society states that changes in the breast can be signal cancer. Signs of breast cancer may include: 1- Breast swelling. 2- Skin irritation or dimpling. 3- Breast pain. 4- Nipple pain or inward turning of the nipple. 5- Redness, scaling, or thickening of the nipple or breast skin. 6- Non-milk nipple discharge. 7- Axillary lumps [37]. These signs and symptoms warrant further evaluation for possible breast cancer. The previous analysis was traditional and performed by pathologists. However, recent advancements have changed this approach. The new direction is focused on using automatic techniques. These techniques are now used to analyze images. [3]. Mammography is another method used for breast cancer diagnosis. It helps in the detection of the presence of tumor tissues. But this traditional method can't detect if these tissues are benign or malignant. So, the most accurate detection method is by taking a biopsy from this mass. Then examine these images under the microscope. These microscopic images are recognized as Histopathology images. There are four categories of breast tissues in histopathology images. The first category is Normal tissue,

Manuscript received [9 January 2025]; revised [13 March 2025]; accepted [10 April 2025]. Date of publication [1 July 2025]. (Corresponding author: Amany Ibrahim).

Amany Ibrahim Department of Computer Science & Engineering, Faculty of Electronic Eng. Menoufia University Menoufia, Egypt.(email: A.ibrahim@el-eng.menofia.edu.eg).

Hanaa Torkey Department of Computer Science & Engineering., Faculty of Electronic Eng. Menoufia University Menoufia, Egypt. (email: htorkey@el-eng.menofia.edu.eg).

Ayman El-Sayed Department of Computer Science & Engineering, Faculty of Electronic Eng. Menoufia University Menoufia, Egypt. (email: ayman.elsayed@el-eng.menofia.edu.eg).



This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see <https://creativecommons.org/licenses/by/4.0/>

which is defined as having no tumors. If the tissue has non-cancerous tumors, then it's considered as benign tissue which is the second category. Cancerous tumors can be classified as the other two categories either in-situ or invasive. The cancer which is not extended beyond the breast lobular system yet is considered as In-Situ cancer. But when the cancer has invaded the breast, it's considered as Invasive cancer shown in fig. (2).

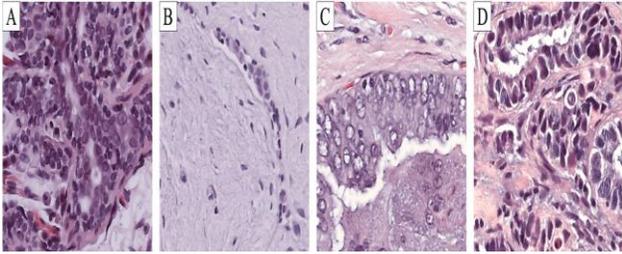


Fig. 2 H&E Histological Images for breast cancer
(A) Benign (B) Normal (C) In-Situ (D) Invasive

Hematoxylin-eosin (H&E) staining has long been rated as the best measure for the histological examination of human tissue. Machine learning (ML) and deep learning (DL) are important fields of investigation for biomedical and bioinformatics researchers. These fields are vital for classifying cancer patients as high risk or low risk. Accurate classification helps guide the right treatment. Early detection of cancer cells is crucial for this process. Timely identification allows for appropriate and effective treatment. Timely identification allows for appropriate and effective treatment. To simulate the advancing in cancer prognosis and treatment, these methods (ML & DL) have been used. CNN has shown great promise in cancer detection and diagnosis in recent studies [4, 5]. Computerized cancer diagnosis from histopathology images had many applied studies. Many times, traditional methods work well. DL does, however, have the benefit of displaying results as though they were human. Deep learning has been used by numerous models to address the segmentation, localization, and classification issues in computer vision [7–12]. In recent years, Graphics Processing Units (GPUs) have revolutionized deep learning, making it faster and more efficient. Unlike traditional Central Processing Units (CPUs), which process tasks step by step, GPUs handle multiple operations at once through parallel computing. This capability significantly accelerates training and inference, making GPUs essential for demanding tasks like image recognition, object detection, and speech processing [6]. With GPU acceleration, deep learning models can quickly analyze large datasets, reduce training times, and improve overall performance. In this study, GPUs are key to managing complex computations, ensuring high-accuracy predictions and seamless real-time processing. Researchers have extensively studied computerized cancer diagnosis using histopathology image analysis. While traditional techniques remain effective in many cases, deep learning methods offer a significant advantage by achieving performance levels comparable to human experts. Various deep learning models have successfully addressed segmentation, localization, and classification challenges in computer vision [7-12]. For this study, we selected the Mask Regional Convolutional Neural

Network (Mask R-CNN) due to its simplicity, flexibility, and generalizability compared to other architectures. An expansion of Faster R-CNN, Mask R-CNN was created especially to solve segmentation problems in computer vision [13]. Interestingly, it did better than any single-model submission in the COCO challenge [13]. Mask R-CNN leverages the regional proposals generated by the Region Proposal Network (RPN) and applies an advanced Region of Interest (ROI) aggregation operation to produce uniform output sizes after feature extraction. Unlike Faster R-CNN, which employs ROI pooling, Mask R-CNN replaces it with ROI-align, ensuring more precise segmentation masks for enhanced instance segmentation. Additionally, the model integrates a Convolutional Neural Network (CNN) as the network head in its architecture. This study proposes a model to address data sparsity in histopathological breast cancer classification by combining multiple high-resolution image patches (2048×1536 pixels) from whole-slide images. A deep CNN is incorporated into the network classification process to extract features efficiently, improving the accuracy of breast cancer diagnosis. We decided to use Mask Regional Convolutional Neural Network (Mask RCNN) design due to its simplicity and flexibility. It also provides a general framework compared to other architectures, each with its own advantages. The Mask RCNN is an extended framework of Faster-RCNN. It solves the segmentation issue in computer vision [13]. In the COCO challenge, it outperformed all single model entries [13]. This model takes advantage of the regional Proposal generated by the Regional Proposal Network (RPN). After feature extraction, and Region of Interest (ROI) aggregation operation is used to generate output with the same size. In the Faster RCNN, ROI pooling has been replaced with ROI-align. This change ensures more accurate segmentation masks for better instance segmentation. Mask-RCNN also has basically CNN which is considered as the network head in the procedure. This research paper proposed a model to solve the data sparsity problem in histological image for breast cancer classification process using combination of several further patches (2048x1536 pixels) from the full Slide available images. Deep CNN is used in network classification to extract features efficiently. This is how the document is structured. Section 2 discusses some related studies. Section 3 will then discuss the dataset and its contents, the technologies used, and the proposed model. Section 4 results of our model and conclusions. Part 5 future work research, and medical imaging, ultimately improving healthcare outcomes and advancing precision medicine.

A. Abbreviations and Acronyms

ML: Machine learning
 DL : Deep Learning
 Hematoxylin-eosin :(H&E)
 RPN: Region Proposal Network
 CNN: Convolutional Neural Network
 GPUs: Graphics Processing Units
 CPUs: Central Processing Units
 ROI: Region of Interest

B. Algorithms

Algorithm 1: MaskRCNN algorithm

Input: Histological images from the BACH dataset

Output: Refined Region of Interest (ROI) proposals

```

1.                                     # Step 1: Load and Preprocess Images
2.     function LoadAndPreprocess(img):
3.         img_resized = Resize(img, (H, W))
4.         img_normalized = Normalize(img_resized)
5.     return img_normalized
6.                                     # Step 2: Generate Anchors and Adjust for Ground Truth Objects
7.     function GenerateAnchors(feature_map, scales, aspect_ratios):
8.         anchors = []
9.         for each scale in scales:
10.            for each ratio in aspect_ratios:
11.                anchor = ComputeAnchor(scale, ratio)
12.                AdjustAnchor(anchor, ground_truth_boxes)
13.                anchors.append(anchor)
14.     return anchors
15.                                     # Step 3: Region Proposal Network (RPN) Activation
16.     function RPN(feature_map, anchors):
17.         proposals = []
18.         for each anchor in anchors:
19.             score, bbox = RPN_ForwardPass(feature_map, anchor)
20.             if score > threshold:
21.                 proposals.append((score, bbox))
22.     return proposals
23.                                     # Step 4: Plot ROI Slices for Visualization
24.     function PlotROIs(img, proposals):
25.         for each proposal in proposals:
26.             DrawBoundingBox(img, proposal)
27.                                     # Step 5: Apply Non-Maximum Suppression (NMS)
28.     function NMS(proposals, IoU_threshold):
29.         proposals_sorted = SortByScore(proposals)
30.         final_proposals = []
31.         while proposals_sorted is not empty:
32.             best_proposal = proposals_sorted.pop(0)
33.             final_proposals.append(best_proposal)
34.             for each proposal in proposals_sorted:
35.                 IoU = ComputeIoU(best_proposal, proposal)
36.                 if IoU > IoU_threshold:
37.                     proposals_sorted.remove(proposal)
38.     return final_proposals
39.                                     # Step 6: Process ROIs for Mask Prediction
40.     function MaskPrediction(final_proposals, feature_map):
41.         masks = []
42.         for each proposal in final_proposals:
43.             mask = MaskHead(feature_map, proposal)
44.             masks.append(mask)
45.     return masks
46.                                     # Main Function to Run the Algorithm
47.     function MaskRCNNPipeline(img):
48.         preprocessed_img = LoadAndPreprocess(img)
49.         feature_map = ExtractFeatures(preprocessed_img, ResNet)
50.         anchors = GenerateAnchors(feature_map, scales, aspect_ratios)
51.         proposals = RPN(feature_map, anchors)
52.         PlotROIs(img, proposals)
53.         refined_proposals = NMS(proposals, IoU_threshold=0.5)
54.         masks = MaskPrediction(refined_proposals, feature_map)
55.     return masks

```

II. RELATED WORK

Research in the early 2000s focused on models to detect abnormalities in mammograms. Traditional approaches often required significant tuning based on the dataset. They produced inconsistent results across different settings. When datasets or parameters changed, these models needed reconfiguration. Such limitations led researchers to explore more adaptive deep learning models for breast cancer segmentation and classification [15,16]. Malarvizhi, A. et al. suggested a Bayesian Interactive Adaboost CNN (B-IAB-CNN) for classification and an Improved-Mask R-CNN (I-MRCNN) for segmentation. While ROI-align addresses quantisation issues to improve feature extraction, I-MRCNN uses a Region Proposal Network (RPN) to improve object detection. B-IAB-CNN improves classification accuracy by combining CNN, Bayesian, and Adaboost techniques.

The BACH dataset, which contains images of normal, benign, invasive carcinoma, and DCIS histology, was used to train the algorithm. 96.32% segmentation and 96% classification accuracy were attained by the model. [17]. Maurya et al. (2024) proposed FCCS-Net, a fully convolutional, attention-based transfer learning model that is designed for breast cancer classification. This model integrates a convolution-driven attention mechanism within a fine-tuned ResNet-18, which helps to highlight key cellular features with using residual connections to maintain learning stability. When tested on the BACH dataset, it demonstrated average classification accuracy of 91.25% and an AUC score of 98.5, precision (91.18), recall (91.38), and an F1-score of (91.28). [18]. Sreelekshmi et al. (2024) proposed a hybrid multi-class classification model that combines CNNs with the Swin Transformer to enhance breast cancer detection and classification. This model combines depth-wise separable convolution for taking localized features with the Swin Transformer's ability to recognize global patterns. They enhanced the accuracy and robustness that were 97.8% classification accuracy, recall (96.8%), precision (97.1%), and an F1-score (97.1%) [19].

A different approach is taken by Kutluer et al. (2019), who combines a unique feature selection method with different deep transfer learning models. Their suggested model obtained a classification accuracy of 92.17% after training on the BACH dataset. [20]. Su et al. (2022) concentrated on mass detection and segmentation. They developed a YOLO and LOGO Transformer-based model. The model trained on the CBIS-DDSM [38] and INBreast [39] datasets. Their approach achieved a 95.7% true positive rate, along with powerful segmentation performance [24]. Bagchi et al. (2022) proposed a deep learning-based classification model that uses a patch-based strategy. Their approach used stain normalization and augmentation to separate high-resolution histopathological

pictures into smaller patches before categorizing them into four groups. A two-stage neural network was used to integrate these patch-level predictions into image-level classifications. On the ICIAR BACH, the model's accuracy was 98.6% for two-class classification and 97.50% for categorical classification [25]. The Deep Ensemble Graph Network (DEGN) was presented by Shwetha G. et al. (2024) for the analysis of breast cancer histology, and it performed flawlessly on the BCSS dataset. DEGN obtained 93% precision, 92% recall, 96% accuracy, and 93% F1-score in Multi-Magnification (MM) mode. It achieved 92% F1-score, 94% precision, 95% recall, and 94% accuracy in Sim (similarity measurement module) mode. [36].

Baroni et al. (2024) proposed a model that used self-attention Vision Transformer (ViT) model. To help in histopathology-based breast cancer classification. Their model use ImageNet pretraining, data augmentation, color normalization, and optimized patch configurations to boost accuracy. Trained on the BACH dataset, the ViT model was later tested on BRACS and AIDPATH to evaluate its generalizability. It achieved classification accuracies of 0.91 on BACH, 0.74 on BRACS, and 0.92 on AIDPATH [35]. Wahab et al. (2019) applied transfer learning and hybrid CNNs to identify mitotic nuclei in histopathology images. Their approach achieved an F-measure of 71.3 [26]. Sunardi et al. (2022) used two deep learning models (CNN and faster RCNN) on mammograph images (MIAS) [40]. They worked to detect the mass in the breast. They found that CNN is applied for mass classification, while Faster R-CNN focuses on mass localization and classification [34]. Rashmi et al. (2023): proposed model for breast nuclei unsupervised segmentation using Applied a multi-channel Chan-Vese. Their approach utilized color segmentation to enhance accuracy, achieving an Intersection over Union (IoU) score of 0.76 and an accuracy of 82% [27].

Samudra et al (2024) presents a novel hybrid approach for semantic segmentation in breast cancer imaging, utilizing a DenseNet-121 backbone integrated with an enhanced Pyramid Scene Parsing Network (PSPNet) and Attention Gate mechanisms. This model addresses challenges in tumor detection accuracy, computational efficiency, and noise reduction. The preprocessing phase employs Adaptive Local Gamma Correction (ALGC) to enhance image contrast, while the Attention Gate highlights critical features, suppressing irrelevant data. Additionally, a Pyramid Dilated Convolution Module (PDM) is introduced to capture global contextual information. The model achieves a significant improvement in predictive accuracy (94.68%) over traditional segmentation models, making it a promising tool for clinical applications in early tumor detection. [33]. All this related work is shown below in table 1.

Table 1: Related work comparison

Ref	Year	Used Dataset	Used Algorithms	Calculated Metrics
[17]	2022	BACH	Improved-Mask R-CNN (I-MRCNN) for segmentation and a Bayesian Interactive Adaboost CNN (B-IAB-CNN) for classification	ACC =96.32% Precision=95 Recall=95.01 F-score=95.69
[18]	2024	BACH	fully convolutional attention-based transfer learning framework	ACC=91.25 recall =91.38 precision =91.18 F-score=91.28
[19]	2024	BACH	combining CNN and Swin Transformer	ACC=97.8 Recall=96.8 Precision=97.1 f-score=97.1
[20]	2023	BACH	Their model forms of recurrent and convolutional deep network	ACC=92.17 recall =92.2 precision =92.4 F-score=92.15
[21]	2020	ICIAR 2018 Breast Cancer Dataset (IBCD)	convolutional neural network (CNN) trained with a new strategy called reversed active learning	ACC= 89.16% to 92.81%
[22]	2022	Used two data sets BCDR-01 and CBIS-DDSM Mammography dataset	Residual U-Net model and use 2 -ResNet for classification	Mean ACC=0.98 Mean IOU=0.94 F-Score=0.98
[23]	2022	DDSM Mammography dataset	Yolo algorithm	ACC = 97.50% F-score=95.23%
[24]	2020	Used tow data sets (CBIS-DDSM and INBreast)	Yolo algorithm	ACC=95.7% IoU=64.0% F1-score=74.5% precision= 65.0%
[25]	2022	BACH dataset	two-stage neural network for patch-based classification	ACC= 97.50% Precision=98.25 Recall= 97.75 f-score= 97.75
[26]	2021	MITOS12 dataset	(CNN) for segmentation, followed by Hybrid-CNN (with Weights Transfer)	Precision=0.506 Recall=0.804 f-score=0.713
[34]	2022	Mammograph dataset((MIAS)	two model CNN and RCNN	ACC for CNN = 91.66% ACC for RCNN = 63.89%
[27]	2023	They used two datasets. Kasturba Medical College (KMC) dataset and BreakHis dataset	Multi-channel Chan-Vese model	ACC= 82.0% IoU= 0.76
[33]	2024	whole-slide images (WSIs)	combines DenseNet-121 as the backbone for feature extraction with a Pyramid Scene Parsing Network (PSPNet)	ACC=94.98% Precision: 94.38% Recall: 94.73% F1 Score: 94.38%
[35]	2024	Bach data set	self-attention Vision Transformer (ViT)	Acc=91% precision= 86% Recall=80% F-Score=85%
[36]	2024	BCSS2021 Dataset BACH Dataset	Deep Ensemble Graph Network (DEGN)	In MM mode: precision =093% recall =92% acc=96% F1-score -93% for SIM mode: precision =94% recall =95% acc = 94% F1-score=92%

III. MATERIAL AND METHOD

A) USED DATASET

The 400 high-resolution H&E-stained photos of breast tissue in the BACH (Breast Cancer Histology dataset) dataset are evenly divided into four groups: benign, invasive, in situ cancer, and normal. Each image is 2048×1536 pixels in size. This dataset, which was developed for the ICIAR 2018 Grand Challenge, supports research on breast cancer segmentation and classification. Expertly labeled by pathologists, it serves as a valuable resource for training machine learning models in automated breast cancer detection, supporting tasks such as classification and semantic segmentation. It's divided to 80% training and 20% validation. It also has separate test dataset contain 100 images also [29]. As shown in table 2.

Table 2 Description of dataset

Class name	Total no. of images	Training images	Validation images
Normal	100	80	20
Benign	100	80	20
In Situ Carcinoma	100	80	20
Invasive Carcinoma	100	80	20
Testing dataset	100		

B) METHODOLOGY

The input images are fed into ResNet-50 and ResNet-101, convolutional networks, which serve as backbone for feature extraction to assess accuracy.

Primarily, these networks identify basic characteristics like edges. By integrating the FPN, high-level features are further retrieved. It aids in the creation of the feature map through the use of ResNet and FPN. PN modules were created to improve multi-scale feature extraction, especially for the analysis of histopathology images.

The proposed model allows to merge low-level fine details with higher-level contextual information to improve overall feature representation. This model is especially useful for precisely detecting cancerous points in histopathological images. The full design of Mask R-CNN is shown in Figure 3.

Our proposed model uses ResNet-50 and ResNet-101 as backbone convolutional neural networks (CNNs) to extract distinguished features from histopathological images to enable accurate detection and segmentation of cancerous regions.

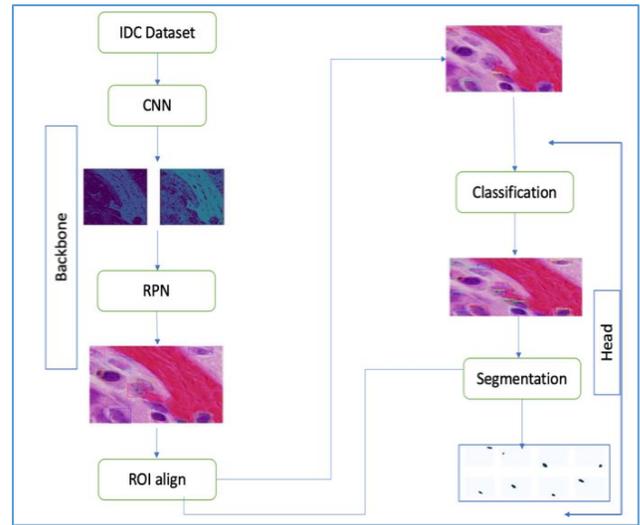


Fig 3 Mask RCNN structure

Firstly, these CNNs detect fundamental structures such as edges and textures, while deeper layers powered by the Feature Pyramid Network (FPN) select more complicated patterns. The FPN enhance multi-scale feature extraction by integrating fine-grained details from lower layers with high-level semantic information, improving the model's ability to identify complex histological structures with greater accuracy.

We apply some preprocessing steps before feeding these images into the CNNs such as normalization (using the COCO dataset's mean and standard deviation), resizing with preserving aspect ratios, and data augmentation techniques like flipping, rotation, scaling, and hue/saturation adjustments. These techniques help differentiate the dataset, minimize the risk of overfitting and improving the model's ability to generalize. However, the in-situ carcinoma class posed challenges in augmentation since its structural inconstancy reduced the effect of transformations like elastic deformations and zooming. The dataset after augmentation is shown in table 3 and augmented version is shown in figure 4. After preprocessing, images pass through ResNet-50 and ResNet-101 for feature extraction within the Mask R-CNN framework.

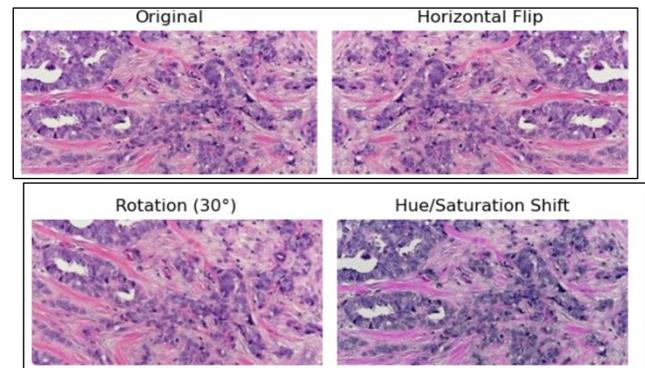


Fig 4 augmented images

Table 3 augmented dataset

Class name	Total no. of images	Training images	Validation images
Normal	400	320	80
Benign	400	320	80
In Situ Carcinoma	202	162	40
Invasive Carcinoma	400	320	80

The extracted features are then processed by the Region Proposal Network (RPN), which scans for potential cancerous regions by proposing bounding boxes based on objects scores. To refine these proposals, the model calculates the Intersection over Union (IoU) between anchor boxes and ground truth objects, ensuring precise region selection.

A crucial component of this process is ROI-Align, which fine-tunes the selected regions by preserving spatial precision through bilinear interpolation. Prior to submitting the features to fully connected and convolutional layers for classification and segmentation, this phase guarantees consistency in feature representation. The segmentation process applies binary masks to the detected regions, providing detailed outlines of cancerous areas. Meanwhile, the classification module, powered by SoftMax regression, categorizes the regions into Normal, Benign, In Situ Carcinoma, or Invasive Carcinoma based on learned patterns. As shown in figure 5.

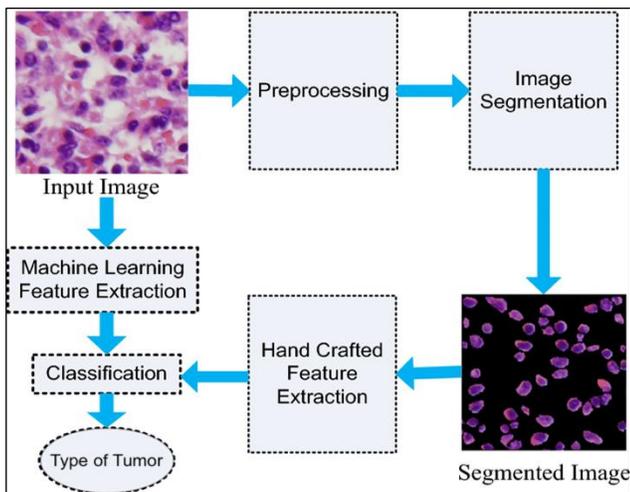


Fig 5 model flow chart

The entire process, from image preprocessing to classification.

The integration of ResNet-50 and ResNet-101 as backbone architectures within the Mask R-CNN framework significantly enhances feature extraction, particularly when combined with the Feature Pyramid Network (FPN). This setup improves the detection and segmentation of intricate structures in histopathological images. Leveraging transfer learning also reduced training time and computational demands while

maintaining high accuracy, underscoring the effectiveness of these architectures in medical imaging applications.

The Feature Pyramid Network (FPN) creates a hierarchical feature representation using a top-down pathway with lateral connections. This allows both high-resolution and low-resolution features to contribute to detecting objects of various sizes. At the same time, the Region Proposal Network (RPN) operates on these enhanced feature maps, scanning them to generate region proposals by predicting bounding boxes and assigning objectness scores.

By leveraging FPN's multi-scale features, RPN improves proposal accuracy, particularly for smaller objects. Instead of working separately, FPN and RPN function together—FPN enhances feature extraction, while RPN uses these refined features to generate more precise region proposals, strengthening the overall object detection framework.

For training, the Adam optimizer was used with a learning rate of 0.001 and a composite loss function that included classification loss, bounding box regression loss, and mask prediction loss. Both ResNet-50 and ResNet-101 are tuned to be more closely aligned with the pre-trained weights to the histological dataset. Training was performed through 50 epochs with mini batches, ensuring stable learning. Fine-tuning pre-trained weights helped reduce computational costs while improving detection accuracy. The dataset was split into 80% for training and 20% for validation, with early stopping applied to forbid overfitting.

Metrics like accuracy (ACC), precision (P), recall (R), and F1-score were computed in order to assess the model's performance. While qualitative evaluations involved superimposing bounding boxes and anticipated masks onto input pictures to confirm segmentation accuracy, confusion matrices offered more in-depth insights into classification errors. Particularly for small and intricate structures in histopathological pictures, object detection was much enhanced by combining ResNet-50 and ResNet-101 with FPN.

By merging, FPN and RPN create a dominant detection system. FPN refines feature extraction, while RPN generates accurate region proposals, ultimately enhancing the overall detection framework.

For implementation, we use Anaconda and Jupyter Notebook, by using PyTorch as the core deep-learning framework. The experiments were performed on a MacBook Air (2020) with an Apple M1 chip, using its integrated GPU for efficient computation.

IV. RESULTS AND DISCUSSION

In this study, we used anchors to assist the RPN identify potential object locations in histopathology images. We implemented a binary classifier that assigns a score based on the likelihood of an object being present in each region to achieve this. The anchors, which have a high Intersection over Union (IoU) score, is called positive anchors and passed on for further processing. The RPN targets were generated by computing anchor grids that covered the entire image at multiple scales. To evaluate the quality of these anchors, we compared their IoU values with ground truth objects. Specifically:

Anchors with $\text{IoU} > 0.7$ were labeled positive, meaning they had a strong overlap with actual objects.

Anchors with IoU < 0.3 were classified as negative and discarded.

Anchors between these values were ignored to avoid adding noise to the training process.

By filtering out irrelevant anchors, we ensured that only the most informative ones were used for training, improving the overall accuracy and effectiveness of the RPN.

To further refine detection, we calculated necessary resizing and displacement adjustments for each anchor to fully cover the ground truth objects. Using the Keras framework, we activated the RPN and visualized region-of-interest (ROI) slices. Additionally, we applied Non-Maximum Suppression (NMS) to remove duplicate detections, streamlining the final output. As shown in algorithm 1.

In the next step, we generated bounding box regression outputs and class probabilities using a classifier proposal. To enhance detection accuracy, we removed background regions around classified nuclei by applying a low-confidence threshold, ensuring that only relevant bounding boxes were kept.

For backbone network analysis, we tested ResNet-50 and ResNet-101, evaluating their performance across all layers and heads. Our results showed a consistent decrease in loss with more iterations, with notable improvements after 50 iterations. This reinforced the effectiveness of our approach in detecting and classifying objects in histopathological images.

The segmentation loss stabilized at this point, reflecting the probabilistic challenges inherent in cancer detection tasks. The results of our classification efforts are summarized in the confusion matrix presented in Figure 6.

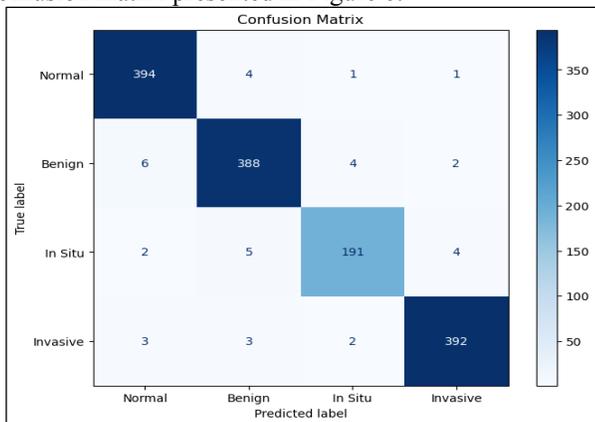


Fig. 6 confusion matrix for Resnet 50

The model reached an accuracy of 97.7% during training and 96.5% on the validation set for ResNet 50, and 96.2% during training and 95.4% for ResNet 101, demonstrating its strong ability to generalize well to new data as shown in Figures 7 and 8.

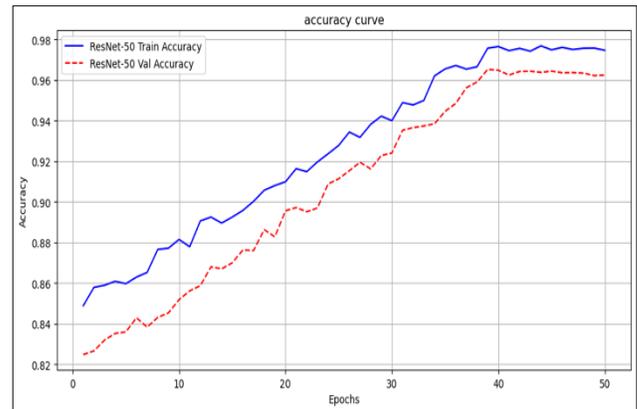


Fig. 7 Resnet 50 Accuracy curve.

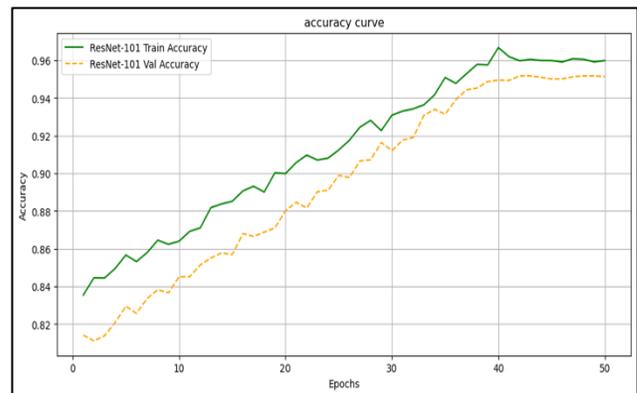


Fig. 8 Resnet 101 Accuracy curve.

The loss function decreased, with the training loss dropping from 1.2 to 0.35, and the validation loss for Resnet 50 and dropping from 1.4 to 0.5 for Resnet 101, indicating efficient learning with minimal overfitting as shown in figure 9 and 10.

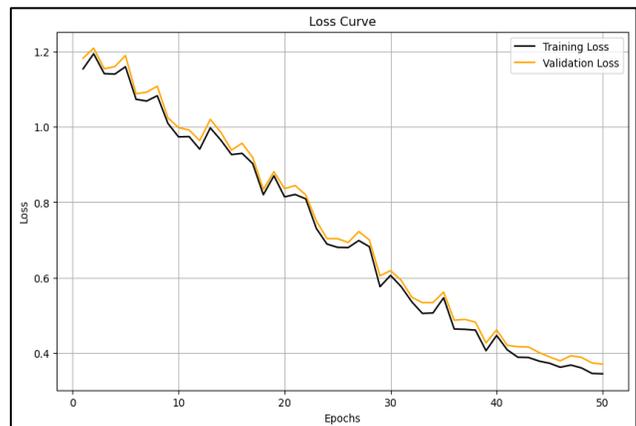


Fig 9 loss curve for resnet50

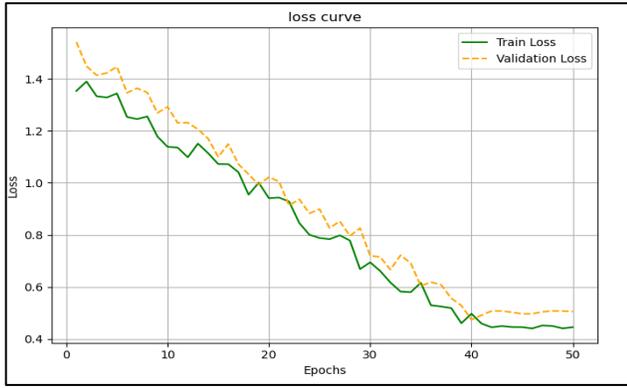


Fig 10 loss curve for resnet101

For the ROC-AUC scores, all classes (Normal, In situ, Benign, Invasive) had values greater than 0.94, with Normal and Invasive achieving the highest AUCs of 0.98 and 0.97, respectively, confirming the model's robust performance across all categories as shown in figure 11.

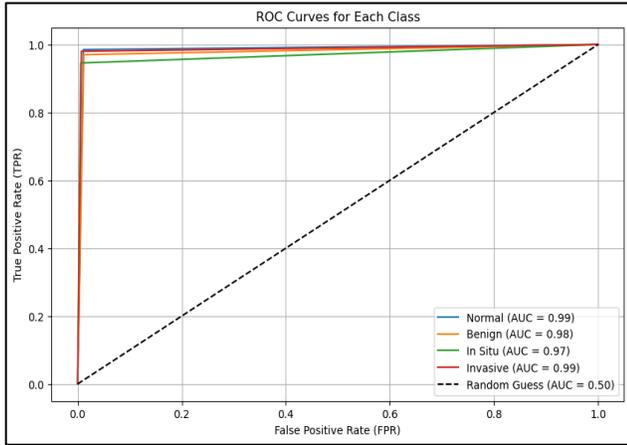


Fig. 11 ROC-AUC curve

These metrics provide a comprehensive evaluation of the model's performance in classifying breast cancer instances.

The RPN targets serve as critical training values, where Precision is defined as the proportion of True Positive (TP) samples relative to the total number of samples anticipated to be positive. There are two potential outcomes for positive predictions: accurately predicting a positive class (TP) or incorrectly classifying a negative instance as positive (False Positive, FP). This relationship can be mathematically expressed as:

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

Recall, on the other hand, is defined as a coverage metric that quantifies the number of true positive samples correctly identified among the actual positive instances. There are also two possible scenarios for recall: correctly predicting the original positive class (TP) or misclassifying a positive instance as negative (False Negative, FN). The calculation for Recall is given by:

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

Through these evaluations, we demonstrate the robustness and reliability of our proposed model in accurately classifying breast cancer instances. The high accuracy and favorable performance metrics underscore the potential of our approach for clinical applications in histopathological diagnosis, thereby contributing to advancements in automated cancer detection methodologies. The first is to make a prediction for the original positive class as the positive class (TP). The second is to make a prediction for the original positive class as the negative class (NP) (FN). The F-measure is a comprehensive evaluation metric.

$$F_1Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

Apparent discrepancies between the precision and recall indices can be avoided. It can combine the results of precision and recall. A high F-meter may indicate a more effective test method. Through these evaluations, we demonstrate the robustness and reliability of our proposed model in accurately classifying breast cancer instances. The high accuracy and favorable performance metrics underscore the potential of our approach for clinical applications in histopathological diagnosis, thereby contributing to advancements in automated cancer detection methodologies. The comparative analysis highlights the superiority of our model over other state-of-the-art methods. As shown in table 4 and figure 12.

Our approach using ResNet 50 and ResNet 101 achieves higher accuracy, precision, recall, and F1 scores than other deep learning models, including self-attention Vision Transformers (ViTs) and Deep Ensemble Graph Networks (DEGN). The improvements are attributed to the efficient feature extraction capabilities of the ResNet backbone combined with the Feature Pyramid Network (FPN) for multi-scale feature learning. Furthermore, the incorporation of Non-Maximum Suppression and optimized anchor selection enhances detection performance, leading to reduced false positives and improved classification accuracy.

Table 4. Comparison between our model and other models. Considering these encouraging outcomes, there are certain drawbacks to our strategy. A significant obstacle is the dataset's class imbalance, which could have an impact on how well the model generalizes to under-represented categories. This could be addressed in future research by using sophisticated data augmentation techniques or by utilizing artificial data generation methods like Generative Adversarial Networks (GANs). Furthermore, while our current approach mostly concentrates on supervised learning, adding self-supervised or semi-supervised learning frameworks could improve model robustness, especially in situations where there is a lack of annotated data. Clinically speaking, the suggested model has a great deal of promise for incorporation into computer-aided diagnostic (CAD) systems, which would help pathologists identify and categorize breast cancer. To guarantee model generalizability in the actual world, additional validation utilizing a variety of datasets from various medical institutes is necessary. Moreover, interpretability remains a critical aspect of AI-driven diagnostics. Future research could explore explainable AI techniques to provide clinicians with insights into the model's

decision-making process, increasing trust and adoption in clinical settings.

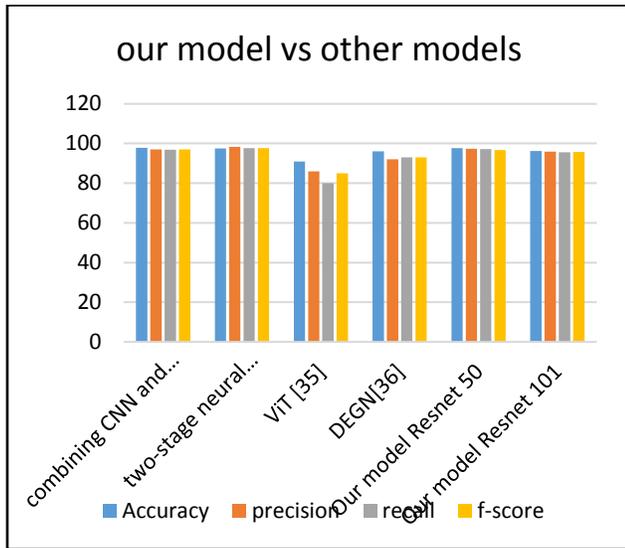


Fig. 12 Comparison between our model and other model

V. Conclusion and future work

The Mask Regional-Convolutional Network (Mask-RCNN) segmentation approach is widely used in contrast to earlier models that did not use segmentation. With a 97.7% classification accuracy for breast cancer tissue pictures, the Mask RCNN with Resnet 50 model outperforms the model without segmentation. when over 40 characteristics are extracted. Because it has less convolutional layers, the epoch (iterative) backbone network ResNet50 performs better than ResNet101 and is computationally less expensive. By including a preprocessing step and expanding the image dataset from 400 to 1402 (477 in each class except In Situ), we decreased the likelihood of overfitting. In the backbone layer, we employed a variety of convolutions and neural network models. The accuracy results were significantly enhanced by this. The accuracy of the results was further increased by altering the number of convolutional layers and using different neural network models in the backbone layer. For instance, the number of convolutional layers in the ResNet50 model, which is currently in use for deep feature extraction, can be altered, as can the hyperparameters for the training and validation datasets. Validation results could be enhanced by adding more filters during training.

We intend to expand our research to more datasets and domains in the future and carry out additional model optimizations throughout the training phase. Other contemporary data augmentation methods (like GANs) ought to be investigated as well. We recognize that a more thorough description of experimental design is necessary to guarantee critical evaluation and reproducibility. We intend to expand our research in subsequent work by evaluating the suggested system's generalizability in practical applications by verifying it on other datasets. We intend to integrate larger and more varied histopathology imaging datasets to further assess the scalability and reliability of our methodology, as the current collection is restricted to 400 samples. This will enable us to

improve the model's functionality and guarantee that it can be used in a range of clinical situations.

Table 4. Comparison between our model and other mode

Trdc l ncdk	@bbt q bx '\$ (Qqdbtrh nm 'S (Qdb 'kk '\$ (E- L d'rt qd '\$ (
bnl alhmf BMM `nc Rv lm Sq mrenq dq	86-7	86-0	85-7	86-0
sv n-rs' fd mlt q'k ndsv ngj Z14\	86-4	87-14	86-6 4	86-64
'UHS(Z24\	80	75	7/	74
'CDFM(Z25\	85	81	82	82
Nt q l ncdk	Qdrm ds 4/	86-6	86-3	86-1
	Qdrm ds 0/ 0	85-1	84-8	84-5

REFERENCES

- [1] World Health Organization, "The Statistics for Breast Cancer," *WHO*. [Online]. Available: <https://www.emro.who.int/media/news/world-cancer-day-2024.html>. [Accessed: Jan. 9, 2025].
- [2] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 2980–2988.
- [3] M. A. Al-Masni *et al.*, "Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system," *Comput. Methods Programs Biomed.*, vol. 157, pp. 85–94, 2018.
- [4] N. Mobark, S. Hamad, and S. Z. Rida, "CoroNet: Deep neural network-based end-to-end training for breast cancer diagnosis," *Appl. Sci.*, vol. 12, no. 14, p. 7080, 2022.
- [5] R. Cao *et al.*, "Prostate cancer detection and segmentation in multiparametric MRI via CNN and conditional random field," in *Proc. IEEE 16th Int. Symp. Biomed. Imaging (ISBI)*, 2019.
- [6] K.-L. Liu *et al.*, "Deep learning to distinguish pancreatic cancer tissue from non-cancerous pancreatic tissue: a retrospective study with cross-racial external validation," *Lancet Digit. Health*, vol. 2, no. 6, pp. e303–e313, 2020.
- [7] H. Gilmore *et al.*, "Accurate and reproducible invasive breast cancer detection in whole-slide images: A deep learning approach for quantifying tumor extent," *Sci. Rep.*, vol. 7, p. 46450, 2017.
- [8] D. C. Cireřan, A. Giusti, L. M. Gambardella, and J. Schmidhuber, "Mitosis detection in breast cancer histology images with deep neural networks," in *Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. (MICCAI)*, 2013.
- [9] A. B. Tosun and C. Gunduz-Demir, "Graph runlength matrices for histopathological image segmentation," *IEEE Trans. Med. Imaging*, vol. 30, no. 3, pp. 721–732, 2010.
- [10] A. Basavanahally *et al.*, "Incorporating domain knowledge for tubule detection in breast histopathology using O'Callaghan neighborhoods," in *Med. Imaging 2011: Comput.-Aided Diagn.*, vol. 7963, *SPIE*, 2011, pp. 305–319.
- [11] T. E. Liang, U. U. Sheikh, and M. N. H. Mohd, "Malaysian car plate localization using region-based convolutional neural network," *Bull. Electr. Eng. Inform.*, vol. 9, no. 1, pp. 411–419, 2020.
- [12] F. A. N. Rashid *et al.*, "Deep convolutional network approach in spike train analysis of physiotherapy movements," in *Adv. Electron. Eng.: Proc. ICCEE 2019*, Springer Singapore, 2020, pp. 159–170.
- [13] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017.

- [14] T. E. Liang, U. U. Sheikh, and M. N. H. Mohd, "Malaysian car plate localization using region-based convolutional neural network," *Bull. Electr. Eng. Inform.*, vol. 9, no. 1, pp. 411–419, 2020.
- [15] J. J. W. Siet et al., "A comprehensive review of tubule formation in histopathology images: Advancement in tubule and tumor detection techniques," *Artif. Intell. Rev.*, vol. 57, no. 1, p. 286, 2024.
- [16] T. Arau et al., "Classification of breast cancer histology images using convolutional neural networks," *PLoS One*, vol. 12, no. 6, e0177544, 2017.
- [17] A. Malarvizhi and A. Nagappan, "Improved Mask R-CNN segmentation and Bayesian interactive AdaBoost CNN classification for breast cancer detection on BACH dataset," *Int. J. Electr. Electron. Res.*, vol. 10, no. 4, pp. 1166–1175, 2022.
- [18] R. Maurya, N. N. Pandey, M. K. Dutta, and M. Karnati, "FCCS-Net: Breast cancer classification using multi-level fully convolutional-channel and spatial attention-based transfer learning approach," *Biomed. Signal Process. Control*, vol. 94, p. 106258, 2024.
- [19] V. Sreelekshmi, K. Pavithran, and J. J. Nair, "SwinCNN: An integrated Swin Transformer and CNN for improved breast cancer grade classification," *IEEE Access*, 2024.
- [20] N. Kutluer, O. A. Solmaz, V. Yamacli, B. Eristi, and H. Eristi, "Classification of breast tumors by using a novel approach based on deep learning methods and feature selection," *Breast Cancer Res. Treat.*, vol. 200, no. 2, pp. 183–192, 2023.
- [21] D. Abdelhafiz, S. Nabavi, R. Ammar, C. Yang, and J. Bi, "Residual deep learning system for mass segmentation and classification in mammography," in *Proc. 10th ACM Int. Conf. Bioinformatics, Comput. Biol. Health Inform. (BCB)*, 2019.
- [22] M. A. Al-Masni et al., "Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system," *Comput. Methods Programs Biomed.*, vol. 157, pp. 85–94, Apr. 2018, doi: 10.1016/j.cmpb.2018.01.017.
- [23] Y. Su, Q. Liu, W. Xie, and P. Hu, "YOLO-LOGO: A transformer-based YOLO segmentation model for breast mass detection and segmentation in digital mammograms," *Comput. Methods Programs Biomed.*, vol. 221, 2022.
- [24] M. A. Al-Antari, S.-M. Han, and T.-S. Kim, "Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms," *Comput. Methods Programs Biomed.*, vol. 196, p. 105584, 2020.
- [25] A. Bagchi, P. Pramanik, and R. Sarkar, "A multi-stage approach to breast cancer classification using histopathology images," *Diagnostics*, vol. 13, no. 1, p. 126, 2022.
- [26] R. Rashmi, K. Prasad, and C. B. K. Udupa, "Multi-channel Chan-Vese model for unsupervised segmentation of nuclei from breast histopathological images," *Comput. Biol. Med.*, vol. 136, p. 104651, 2021.
- [27] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 2980–2988.
- [28] Kaggle, "BACH: Breast Cancer Histology Images," [Online]. Available: <https://www.kaggle.com/datasets/truthisneverlinear/bach-breast-cancer-histology-images/data>. [Accessed: 09-Jan-2025].
- [29] N. Wahab, A. Khan, and Y. S. Lee, "Transfer learning based deep CNN for segmentation and detection of mitoses in breast cancer histopathological images," *Microscopy*, vol. 68, pp. 216–233, 2019.
- [30] R. Rashmi, K. Prasad, and C. B. K. Udupa, "Multi-channel Chan-Vese model for unsupervised segmentation of nuclei from breast histopathological images," *Comput. Biol. Med.*, vol. 136, p. 104651, 2021.
- [31] N. Sindhwan et al., "Performance analysis of deep neural networks using computer vision," *EAI Endorsed Trans. Ind. Netw. Intell. Syst.*, vol. 8, no. 29, p. e3, 2021.
- [32] S. Samudrala and C. K. Mohan, "Semantic segmentation of breast cancer images using DenseNet with proposed PSPNet," *Multimedia Tools Appl.*, vol. 83, no. 15, pp. 46037–46063, 2024.
- [33] I. Gupta et al., "A deep learning based approach to detect IDC in histopathology images," *Multimedia Tools Appl.*, vol. 81, no. 25, pp. 36309–36330, 2022.
- [34] S. Sunardi, A. Yudhana, and A. R. W. Putri, "Mass classification of breast cancer using CNN and faster R-CNN model comparison," *Kinetik: Game Technol., Inf. Syst., Comput. Netw., Comput., Electron., Control*, 2022, pp. 243–250.
- [35] G. L. Baroni et al., "Optimizing vision transformers for histopathology: Pretraining and normalization in breast cancer classification," *J. Imaging*, vol. 10, no. 5, p. 108, 2024.
- [36] S. G. Krishnappa and K. R. Udaya Kumar Reddy, "Enhancing histopathology breast cancer detection and classification with the deep ensemble graph network," *SN Comput. Sci.*, vol. 5, no. 5, pp. 1–11, 2024.
- [37] American Cancer Society, "Breast cancer signs and symptoms," [Online]. Available: <https://www.cancer.org/cancer/types/breast-cancer/screening-tests-and-early-detection/breast-cancer-signs-and-symptoms.html>. [Accessed: 09-Jan-2025].
- [38] M. Amgad et al., "Structured crowdsourcing enables convolutional segmentation of histology images," *Bioinformatics*, vol. 35, no. 18, pp. 3461–3467, 2019.
- [39] M.-L. Huang and T.-Y. Lin, "Dataset of breast mammography images with masses," *Data Brief*, vol. 31, p. 105928, 2020.
- [40] S. Charan, M. J. Khan, and K. Khurshid, "Breast cancer detection in mammograms using convolutional neural network," in *Proc. 2018 Int. Conf. Comput., Math. Eng. Technol. (iCoMET)*, pp. 1–5, IEEE, 2018.