



Enhancing Breast Cancer Diagnosis: A Semi-Supervised Deep Learning Model for Automated MRI Tumor Segmentation

Rayan A. Alsemmeari¹, Zakir Iqbal², Sheikh Tahir Bakhsh^{3*}, Bandar M. Alghamdi¹,
Abdulaziz A. Alsulami⁴, Badraddin Alturki¹

¹ Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

² Department of Computer Science, COMSATS University Islamabad, Abbottabad 22060, Pakistan

³ Cardiff School of Technology, Cardiff Metropolitan University, Cardiff CF5 2YB, UK

⁴ Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

Corresponding Author Email: sbakhsh@cardiffmet.ac.uk

Copyright: ©2025 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ts.420212>

ABSTRACT

Received: 29 September 2024

Revised: 21 January 2025

Accepted: 18 March 2025

Available online: 30 April 2025

Keywords:

automated medical imaging, Segment Anything Model (SAM), medical image analysis, tumor detection, annotated MRI scans

The upsurge in the number of breast cancer patients makes it important to diagnose the disease early so that effective treatment can save the patient's life. Breast cancer diagnosis is challenging; however, the adoption of various deep-learning techniques has made this hard work much more accessible for radiologists to diagnose breast cancer at an early stage. Numerous applications have been developed to provide practical solutions to aid the radiologist in medical image analysis. Magnetic Resonance Imaging (MRI) is considered the most accurate screening technique for breast cancer identification, and it has substantially contributed to decreasing the mortality rate by early breast cancer detection. In MRI imaging, breast tumor detection and segmentation are still regarded as a critical task due to the limited availability of annotated MRI scans that require more laborious to annotate data with accurate ground truth which is a time-consuming process and less feasible in medical imaging. This research presents a Semi-Supervised Deep Learning for Automated Tumor Segmentation (SDATS). A semi-supervised learning model is employed, utilizing labeled and unlabeled data for segmentation. Segment Anything Model (SAM) is used for tumors, using bounding boxes to isolate regions of interest and generating precise segmentation masks. The YOLOv8 model is utilized for breast tumor detection, identifying bounding boxes for regions of interest. Integrating YOLOv8 and SAM makes the proposed model more rigorous and aims to enhance efficiency without using pixel-level annotation for segmentation. This allows for more efficient processing of large datasets and accelerates the diagnostic process. Furthermore, the SDATS diagnoses breast cancer with quicker and more precise automatic segmentation.

1. INTRODUCTION

To cure breast cancer on time and keep away from numerous useless tests, proper diagnosis techniques for the detection of breast cancer diagnosis is very crucial. Recently published data by the World Health Organization (WHO) suggested that breast cancer is responsible for 14% of cancer-associated deaths among women and 23% of all cancer cases worldwide [1]. Breast cancer is the top-ranked cancer basis of death for women after lung cancer. A substantial statistical rise is expected in breast cancer cases by 2040 [2]. Late breast cancer treatments result in harmful cancer stages and, as a result, a lower survival rate [3]. Thus, early cancer detection could significantly reduce the mortality rate [4]. Breast cancer is typically triggered by the uncontrolled development of abnormal cells that arise from the inner milk lobules or ducts. Tumor and microcalcification are two main types of breast cancer, which can be cancerous or noncancerous cells [5].

Early breast cancer detection is critical for the well-being and survival of patients [6]. The abnormal growth of cells causes cancer tumors, which attack all the tissues surrounding the human body. Cancer can be started everywhere in the body, composed of a vast number of cells. Normally through cell division human cells develop and multiply continuously to create new cells as the organism requires them. New cells adjust themselves in the location of aged and damaged cells but when the orderly process collapses, the abnormal cells grow and multiply. These abnormal tumors are formed by cells and are the knobs of tissue.

Cancerous tumors are also called malignant tumors. They spread into, or attack nearby tissues in the body and can move to distant places to form new tumors that process is called metastasis. Non-cancerous tumors, also called benign cancers don't invade, or attack neighboring tissues. When these tumors are removed, they typically don't grow back while cancerous tumors occasionally grow back. Benign growths can be

relatively large. Several can cause severe indications or be life-threatening like benign tumors in the brain. Breast cancer has been the leading cause of death among women in recent years due to a lack of appropriate diagnosis tools. Figure 1 shows according to the WHO Annual Report 2020, about 0.5 million fatalities are reported worldwide due to breast cancer.

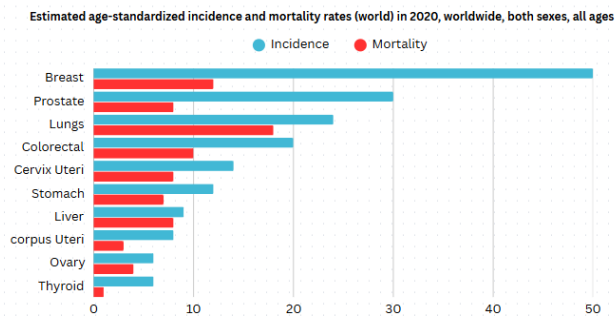


Figure 1. Estimated age standardized incidence and mortality rate [7]

Machine learning plays a vital role in the control of the death rate of breast cancer patients. Using machine learning approaches various tests for the detection of breast cancers have been used so far, for example, Mammograms, Ultrasound, MRI, and biopsy. Magnetic Resonance Imaging (MRI) stands out for breast cancer screening modality at the top level, and this MR imaging is different from the traditional way, with no x-ray experience and no breast compression under controlled pressure [8]. Instead, this method uses big magnets, magnetic fields, and radio waves to make complete breast images. The images cause the breast to have no movements with the women lying down being made depending on the vital energy magnetic field and radio-frequency radiation to make the breast tissue. The breast images are detailed and clear, which indicates that it is possible to detect any small or early changes that signal cancer or review the cancer tumor in the dense breast tissue [9]. The final image is taken by the doctor for MRI benefits and risks analysis. Computer Aided Diagnostics (CAD) approaches which are being developed to allow radiologists to locate and detect abnormal cells in the breast on MRIs, may enhance accuracy while simultaneously limiting the number of potential victims. The feasible combination of computer-aided technology with MRI scans to detect and diagnose breast cancer will, in turn, contribute to patient success [10].

Eliminating pixel-level annotation in breast cancer detection and segmentation research is critical for several reasons. It reduces annotation time by eliminating the necessity of pixel-by-pixel labeling of individual images [11], allowing researchers to concentrate on model creation. It enhances scalability and generalization using a semi-supervised or weakly supervised approach, which will work by combining labeled and unlabeled data. This method is more clinically realistic, as it is more comparable to how radiologists operate in practice. A radiologist does not label every pixel, but regions of interest according to contextual signals are made. Besides, collecting fully annotated datasets is expensive. Ultimately, the ignoring of the pixel-level annotation makes it possible for models to deal with the wide variety of breast tumor forms and textures and achieve much higher segmentation accuracy. The process of segmenting

breast cancer tumors in MRI imaging is crucial to the anticipated diagnosis and planning of treatment. Nevertheless, the pixel-level annotation method used in the analysis described several constraints, this method is expensive and not scalable because of the consumption of time.

Breast cancer continues to be a major global health burden that impacts millions of women and their families each year. The identification of early manifestation and precise diagnosis is highly essential for the effective treatment of patients. The manual segmentation of tumors from medical images is tedious, time-consuming, and error-prone. The literature review reveals that many techniques have been used in the breast cancer tumor detection and segmentation task in the Magnetic Resonance images using a supervised learning approach. However, these techniques fail to generalize, because of the extremely limited dataset size. Currently, the datasets are manually annotated which is a cumbersome, time-consuming process and requires expert knowledge to accurately outline the ground truth images. Therefore, there is a need for automated systems that can auto-generate accurate segmentation data annotations for breast cancer tumors.

The proposed *SDATS* deals with the issue of unbalanced techniques based on feature extraction by proposing this innovative semi-supervised learning framework for the tasks of breast cancer detection and tumor segmentation. However, for this particular effort, the proposed model focuses specifically on the automation of segmentation dataset generation. The system reduces the costs of annotation, given the huge amounts of unlabeled data on one hand and the small number of annotated expert images on the other hand. The research proposes a machine-assisted deep learning labeling technique for generating a segmentation dataset using a detection model with a SAM without doing pixel-level annotation in breast cancer MR imaging. The SAM is an instant segmentation model that can provide a superior quality mask to the entire image or targeted object present in the image using the prompts method [12], but it lacks producing labels [13]. This research solves the limitation using bounding box prompts from the detection model to pass the SAM model to produce a high-quality mask on the object of interest (tumor) with corresponding coordinates. The burden of annotation is significantly reduced by avoiding pixel-level annotation which allows for easy development. Instead of relying on a substantial number of manual annotations, the addition of enough unlabeled data will improve the weaknesses. Since unlabeled data is employed, the small, labeled dataset's limitations are eliminated. The research aims to connect advanced technologies simultaneously with practical applications and facilitate breast cancer diagnoses through efficient automated image analysis.

The remainder of this paper is organized as follows: Section 2 reviews the existing literature review along with an analysis of related work providing context and highlighting gaps that this research aims to fill. Section 3 details the proposed *SDATS* methodology, including the integration of a machine-assisted deep-learning labeling technique that leverages a detection model and the SAM to generate high-quality masks without pixel-level annotation. Section 4 presents the results, demonstrating the effectiveness of the approach in reducing the annotation burden while improving segmentation accuracy. Finally, Section 5 concludes the paper, discussing the implications of the findings and suggesting directions for future research.

2. REVIEW OF DEEP LEARNING TECHNIQUES AND CURRENT APPROACHES IN BREAST CANCER TUMOR SEGMENTATION

Breast cancer is one of the most common and harmful concerns of women population in many countries, particularly in Western countries. About 11% of women are suffering from breast cancer. Different organizations have published several reports about the severity of breast cancer according to several reports it is considered the second most common cancer type after lung cancer and the fifth most common cause of death. World Health Organization has presented that over 400,000 female deaths are due to breast cancer every year. American National Cancer Institute reported that in every 3 minutes, one female is diagnosed with breast cancer, and in every 13 minutes, one female is killed by that disease.

To establish a foundation for the proposed research, it is essential to explore the advancements and challenges in the field of breast cancer diagnosis, particularly in the context of image analysis and segmentation. Traditional computer-aided diagnosis (CAD) systems have long been employed for breast cancer detection, but they often fall short due to limitations in diagnostic accuracy and the complexity of accurately identifying tumor characteristics. Given these challenges, the integration of deep learning approaches offers a promising alternative, addressing many of the shortcomings inherent in conventional methods. Deep learning approaches have a great deal of interest and consideration in the area of image recognition, segmentation, object detection, and computer vision in recent years [14]. Deep learning has been used to address the shortcomings of the conventional [15], this is because of limited diagnostic accuracy. Also, the extreme variance in their form, texture, color, and size makes the diagnosis procedure difficult for conventional techniques, as well as the high degree of similarities in breast tumors [16]. Using the advantages of a deep learning CAD system can present their capabilities to learn complex deep features to improve the diagnosis performance [15]. Even though several systems have used conventional machine-learning techniques for breast cancer segmentation, the usage of deep learning detection and segmentation has been limited in the field of breast cancer MR imaging.

The study by Benjelloun et al. [17] emphasizes the use of machine learning for breast cancer tumor segmentation in (Dynamic Contrast-Enhanced MR Imaging), which is a vital move in breast cancer detection and treatment. To address these issues and the limitations of data used for training in medical imaging, the research utilizes a deep neural network built on the U-Net architecture, that has proven effective in image segmentation projects. The U-Net architecture, which is an enhancement of the fully Convolutional Neural Network, comprises standard FCNN layers coupled with up-sampling layers to increase data size, and it integrates previous feature maps into the final ones to improve image representation learning. Zhang et al. [18] discussed, that breast cancer is a major global health risk for women. This work uses artificial intelligence, namely deep learning, to improve breast cancer tumor identification and categorization. The study uses a kind of convolutional neural network known as Mask R-CNN to discover and identify possible lesions in breast MRI data. The limitations of earlier computer-aided diagnosis systems, which relied on labor-intensive human procedures with widely varied accuracy and preset features, are intended to be addressed by this novel methodology. The work successfully evaluates

region proposals, recognizes objects, and completes segmentation all at once by modifying the Mask R-CNN architecture. Further actions may be taken if the precise site of the tumor has been identified. Praveena and Kumar [19] have employed a deep learning approach that can be used to provide enhanced detection aid for breast cancer. The basic idea behind the technique is to build a machine-learning model to discover the problem by analyzing breast cancer MRI data. DMFTNet model was employed for breast cancer detection with semantic segmentation, manual annotation approach was adopted to label the data.

The research conducted by Guo et al. [20] have developed a technique for detecting and segmenting breast cancer tumors by employing a Convolutional Neural Network with a support vector machine (SVM). The model was trained on a cloistered dataset having 272 patients MRI images from the affiliated hospital of Fujian Medical University. Data was labeled manually by a radiologist. This model did not achieve outstanding results when compared with other models. Xu et al. [21] they proposed a framework of automated deep learning for MRI utilizing the U-Net model. A private dataset was used having 301 patients specifically confirmed biopsy stages III with 744 MRI images. Focusing on breast cancer detection and treatment, Park et al. [22] proposed a deep learning model for three-dimensional tumors of breast cancer segmentation in MRI modality. U-Net transformer UNETER was utilized for automatic segmentation tasks. The Bounding box is used for weak annotation as a ground truth for tumor segmentation to train the model on the dataset of DCE-MRI having 736 images. Yang et al. [23] presented a multi-class semantic segmentation of breast tissue using a Haar wavelet pooling model based on U-Net. In the labeling dataset noise was also removed using the Otsu threshold with a median filter from the MRI images. The dataset is taken from The Cancer Imaging Archive (TCIA). To avoid manual annotation, template-based segmentation is utilized which is less effective in the medical imaging Field.

Swiderski et al. [24] investigated a method for breast cancer diagnosis and tumor detection using an autoencoder-based generative adversarial network (AGAN), a modified GAN architecture. CNN has been used to categorize breast cancers, and AGAN is used for data augmentation. Image identification, similarity detection tasks, and machine learning approaches like deep convolutional neural networks have surpassed cutting-edge visual recognition jobs. In the study by Kim et al. [25], breast abrasion was detected as clotted cell clusters that develop into tumor cells. In this regard, an efficient and improved breast tumor detection method has been outlined by using MRI images that not only offer more rapid detection but also provide competent accuracy as compared to other available works. Various breast abrasion areas surrounded by real breast tumors are not breast tumors but give out several issues, therefore their detection and identification come to be most challenging. To overcome these issues GAN model with the breast tumor local histogram treatment was integrated. In the present scenario, a mathematical morphological approach was incorporated rather than using conventional techniques like filtering approaches in which shape and size features were used for identification.

Existing literature listed in Table 1 and reviews have shown that many approaches have been used to detect and segment breast tumors from MR images. However, these methods are used for very small data sizes because of the limited availability of annotated MRI scans. Most existing work suffers from manual annotation to train the model on a small

number of images to segment tumors. It has been identified that most of the literature reviewed showed how well artificial intelligence solves the problem of breast cancer detection in the field of MRI. Breast cancer diagnosis is a sequential process, which is why developing a step-by-step strategy is critical. Specifically, the research investigates how deep learning can be applied to breast cancer diagnosis and

prognosis using imaging modalities such as MRI, mammograms, ultrasound, and biopsy samples. The goal is to leverage deep learning techniques to enhance the accuracy of tumor detection and segmentation in breast cancer diagnosis. Obtaining data on both traditional and computational procedures.

Table 1. Analysis of related work section and comparison

Author	Year	Model and Input Dataset	Method Type	Performance Evaluation	Key Findings	Weakness
Benjelloun et al. [17]	2018	Model: U-Net. Dataset: private Dataset of 301 Patients	Tumor Segmentation	Intersection Over Union (IoU)	76.14%	Small data size is used to train the model, U-Net Models require a large dataset to produce authentic results.
Hamad et al. [26]	2018	Model: FCM and Thresholding. Dataset: Private Dataset containing 75 images	Mass segmentation	AUC	90%	The model may not be generalized and scalable due to the small data size used to train the model.
El Adoui et al. [27]	2019	Model: U-Net and Seg-Net. Dataset: 86 DCE-MRI of 43 Cancerous patients. Reference: Manual Annotations	Tumor Segmentation	Mean Intersection Over Union (IoU).	Seg-Net: 68.88%, U-Net: 76.14%	Small data size is used, training U-Net models efficiently requires a sizable quantity of annotated data, and generating accurate ground truth annotations for medical images is a time-consuming procedure, also requires domain expertise.
Li et al. [28]	2019	Model: U-Net. Dataset: 313 Patients MRI images. Reference: Manual Annotation.	Mass segmentation	Dice Similarity Coefficient (DSC)	77.6%	U-Net models are sensitive to image quality and do not perform well when noise is present in the image.
Guo et al. [20]	2022	Model: CNN and SVM. Dataset: Private Dataset of 272 Patients of MR Images.	Tumor Segmentation	Dice Similarity Coefficient (DSC)	93%	In small data sizes, SVM does not perform well when the dataset has more noise, the proposed model may not be able to detect various small tumors that are present in hard-to-reach areas of the breast.
Haq et al. [29]	2022	Model: cGAN. Dataset: RIDER from TCIA of 500 MRI images	Tumor Segmentation	Dice Similarity Coefficient (DSC), Intersection Over Union (IoU)	DSC: 85% IoU: 77%	Getting noticeable outcomes from cGAN requires large and varying quantities of annotated data, which is expensive in the medical sector.
Zhang et al. [18]	2022	Model: Mask R-CNN. Dataset: From Private Hospital of 339 MR images	Lesion Segmentation	Dice Similarity Coefficient (DSC)	0.79%	Only 339 images are used to train the model, due to the two-stage architecture Mask R-CNN is more complex to train and relatively slow.
Praveena and Kumar [19]	2023	Model: CNN, DMFTNet. Dataset: From Private Institute of 400 Patients	Tumor Segmentation	Dice Similarity Coefficient (DSC)	76.14%	When data is limited CNN models may occur with an overfitting problem.
Park et al. [22]	2024	Model: U-Net. Dataset: DCE-MRI of 736 images	Lesion Segmentation	Dice Similarity Coefficient (DSC)	0.75	The proposed model struggled to capture small size of lesions in MR images.
Yang et al. [23]	2023	Model: U-Net. Dataset: Taken from TCIA. Template-based segmentation	Tissue Segmentation	Mean Intersection Over Union (mIoU)	87.48%	U-Net model is expensive because it requires a lot of annotated data to train the model.

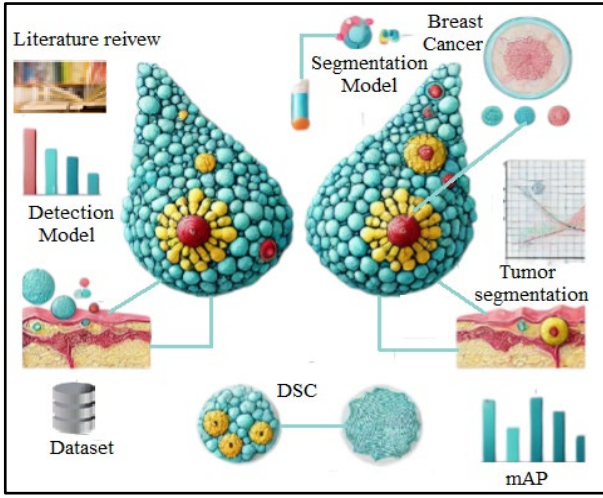


Figure 2. MRI-based breast cancer detection and tumor segmentation

The main issue that has been identified after reviewing the literature is the availability of annotated data due to the very time-consuming and labor-intensive process and the need for relevant field knowledge to exactly mark the boundary of cancer. By solving this problem, deep learning has attained a lot of success in the medical field as illustrated in Figure 2.

Recently deep learning techniques have been wrought in different fields like breast cancer diagnosis based on different types of imaging modalities. Deep learning achieved excellent performance as compared to other traditional machine learning techniques. This study aims to improve existing manual labeling techniques for breast cancer segmentation datasets by integrating state-of-the-art detection models with general-purpose segmentation models.

In this paper, we propose an SDATS model, which incorporates the YOLOv8 detection model with SAM to detect tumors in breast cancer and segment breast cancer tumors. However, in comparison to classic segmentation methods, like Mask R-CNN UNet or DMFTNet which are dependent on pixel-level annotations, SDATS drastically decreases the annotation hassle through the use of bounding box prompts provided by YOLOv8 and subsequent prompt-supported segmentation via SAM.

Leveraging the best in real-time detection (YOLOv8) and efficient segmentation (SAM), SDATS can yield state-of-the-art segmentation precision with lower reliance on manual annotation, making it more scalable and computationally efficient. Unlike Mask R-CNN, which is a two-stage process, we combine detection and segmentation in a single pipeline, improving both speed and accuracy. Moreover, U-Net and DMFTNet usually need extensive annotated datasets. At the same time, SDATS utilizes semi-supervised learning to make the most out of both labeled and unlabeled data, thus mitigating the issue of scarce annotated data.

3. THE SEMI-SUPERVISED DEEP LEARNING-BASED BREAST CANCER DETECTION AND SEGMENTATION IN MRI IMAGES

In the proposed SDATS model, Magnetic Resonance (MR) images are grasped as input dataset American College of Radiology Imaging Network (ACRIN), to accurately detect,

segment, and auto-generate annotation for the segmentation dataset of breast cancer tumors within MR images. Through the You Only Look Once (YOLO) model the tumors are precisely detected from the image. By integrating the YOLO detection model with the SAM, an automated annotation system is developed for breast cancer tumor segmentation. This approach effectively addresses the challenges of time-consuming and costly manual annotation, providing a more efficient and scalable solution for creating high-quality segmentation datasets. The proposed SDATS model segments tumor areas from the rest of the image, which makes statistical data analyses possible. This can be used by researchers and clinicians for diagnosis and treatment planning. The proposed SDATS model with the combination of YOLOv8, auto annotations, and the Segment Anything Model (SAM) can provide robustness and generalizability for tumor localization and segmentation. A solution that can serve to improve clinical decision-making through medical image analysis.

To quantify the reduction in annotation effort, we conducted a comparative analysis. In practice, manual pixel-level annotation required around 5 minutes/image on average, whereas the proposed SDATS model instrumented SAM's bounding box prompts to reduce the time requirement to 30 seconds/image. This means a 90% annotation time reduction. In addition, the approach generalizes well to larger datasets. Using SDATS, the annotation time for a dataset containing 15,000 images was ~125 hours as opposed to 1,250 hours for the standard manual annotation approach.

Figure 3 shows the entire pipeline from the input to output in detail. The first step involves feeding an input of raw medical image i.e., MRI scan for which the model needs to indulge in the segmentation process. Semi-supervised machine-assisted labeling workflow for tumor detection and segmentation in breast cancer MRI. MRI images are likely used for medical diagnostics and deliver detailed, high-definition images of soft tissues that can be crucial in identifying abnormalities (like tumors) in the breast. These images are utilized to detect and segment as the basis for the model automatically. The input to the model is an MRI image which is passed through the YOLO deep learning detection model. The YOLO is a popular single-step object detection algorithm that goes through the image only once and finds areas in which tumors might be located. One of the key capabilities of YOLO is its real-time object detection, which enables us to effectively detect sensible regions or Regions of Interest (ROIs) within MRI scans. It is worth noting that the model does not just classify these regions, it also draws boxes around detected areas which provides a visual cue, showing roundabouts where the supposed tumors can be found. Also, YOLO supplies class probabilities with every box, as an estimate of how likely each localized region corresponds to a breast tumor. This step is important since it helps us to distinguish whether a region is non-cancerous from cancerous regions or possibly noise and directs the segmentation work only to which is more likely a tumor location.

$$\text{Let } I \in \mathbb{R}^{H \times W \times C} \text{ represent the input MRI image} \quad (1)$$

where:

- H is the image's height.
- W represents the image's width, and

C represents the number of channels ($C = 1$ for grayscale MRI).

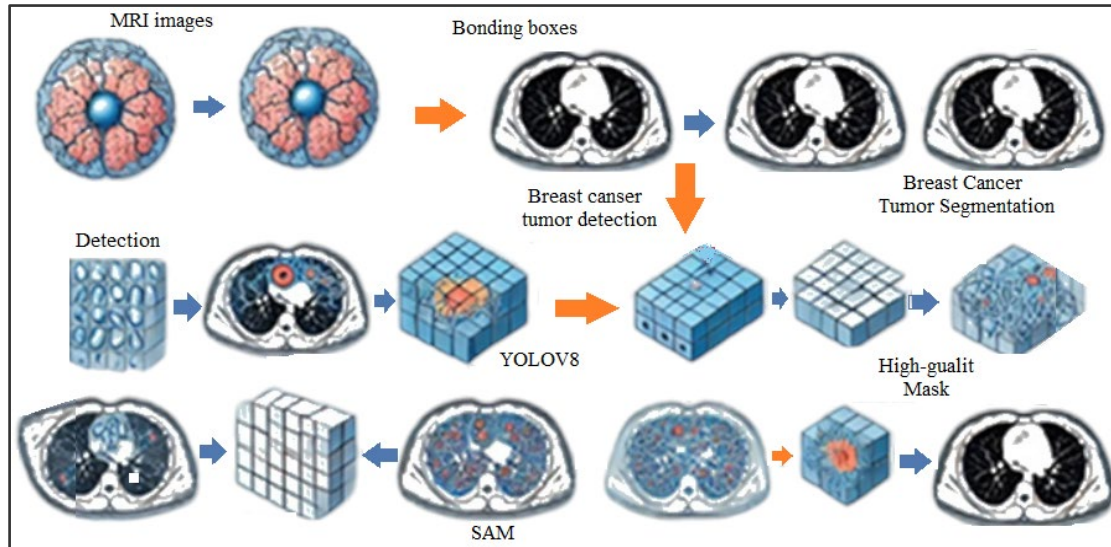


Figure 3. Semi-supervised labeling for breast cancer detection in MRI

Once the ROI detection and bounding box generation have finished in the first stage of tumor localization based on which the detection model returns the possible bounding boxes together with their associated class probabilities. The meritorious boxes indicate a very coarse localization of the prospective area where tumors are located, which makes the deep learning-based detection algorithms specifically designed to avoid misrecognizing normal structures more effective in that focus. Though most areas containing tumors are likely to be encompassed by the bounding boxes, this results in reduced unnecessary segmentation effort not only on irrelevant areas of the MRI images. Which increases efficiency and makes the system more accurate. These bounding boxes are called forward into the next part of the model so that the other stage in modeling will not have to analyze this and can focus on segmentation. In the first progression of studies, the SAM is used, to produce segmentation results. Containing up-to-the-minute segmentation models and a highly accurate SAM model, it receives a pair of bounding boxes, which act as prompts, and an MRI image. MRI image is encoded by an Image Encoder which gives a learned representation of the raw image data, this embedding shows enough resolution to discriminate between images but is not useful for direct transformation into a segmentation mask, this embedding is a compact representation of the image, containing information about color and spatial layout necessary for segmentation. The image is encoded into a low-dimensional space that allows SAM to gracefully process the image by keeping essential information required for precise drawing of tumor boundaries.

Meanwhile, the prompt encoder deals with the bounding box and helps to let segmentation work on YOLO-detected region of interest (ROI). The prompt encoder ensures the relevant parts of the image (i.e., where a potential tumor has been detected) that should be considered during the segmentation process. By integrating the image encoder and prompt encoder, SAM learns to produce more informative and discriminative segmentation results which can effectively suppress false positives when interpreting non-tumor regions. The last step of the process is the lightweight mask decoder, this step produces as a final output the fine segmentation mask. The image and the bounding box prompt are generated by combining information from both. This mask corresponds to the different regions of a tumor in an image with a pixel-level

clear boundary that can be applied for downstream clinical exploration or dataset preparation. A lightweight decoder is used, making the Mask generation process computationally inexpensive and thus delivering segmentation without an accuracy trade-off. This method substantially reduced the expensive and time-consuming manual annotations required for tumor detection and segmentation in medical imaging. By combining YOLO for detection, and SAM for segmentation, the system can efficiently deal with large amounts of MRI data to generate high-quality segmentation masks that could be useful for diagnosis, treatment planning, or as part of a labeled dataset to form the basis of further research and development.

3.1 The proposed SDATS model utilizing YOLOv8 architecture for breast tumor detection

The decision to use the YOLOv8 model for breast cancer diagnosis in this research is based on its advanced capabilities in object detection and classification, making it highly effective for analyzing medical images such as mammograms, MRIs, and ultrasound scans. It is the most famous within the Computer Vision (CV) community, the latest version keeps the YOLO legacy alive and helps in getting state-of-the-art results for image or video analytics with an easy-to-implement framework. YOLO is an object detection algorithm introduced in a 2015 research paper by Joseph Redmon and Ali Farhadi. The architecture of YOLO is a huge step forward in the realm of real-time object detection, outperforming its predecessor: The region-based Convolutional Neural Network (R-CNN). YOLO is a single-shot detector which means unlike its predecessors (SSD, RetinaNet) there are multiple passes involved before the final prediction can be made. It uses a neural network to predict bounding boxes and class possibilities unswervingly. The YOLO model divides the input image I into a grid of $S \times S$ cells. Each cell is responsible for predicting several bounding boxes and associated confidence scores for whether the box contains a tumor.

For each cell, YOLO predicts:

B bounding boxes,

P_c : Confidence score for detection of an object (tumor), and

Parameters (x, y, w, h) : the bounding box's coordinates and measurements.

Therefore, for every cell i , its output is:

$$\text{YOLO Output}_i = \{(x_i, y_i, w_i, h_i, P_{c,i})\} \quad (2)$$

where:

- Confidence that the bounding box wraps around a tumor $P_{c,i} \in [0,1]$
- (x_i, y_i) are the bounding box center's coordinates,
- It takes two parameters, w_i and h_i which are the width and height of the bounding box respectively.

It must minimize the following cost function that incorporates localization, confidence, and classification errors:

$$\begin{aligned} L_{YOLO} = & \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{i,j}^{obj} ((x - \hat{x})^2 + (y - \hat{y})^2 + (w \\ & - \hat{w})^2 + (h - \hat{h})^2) \\ & + \lambda_{\text{coof}} \sum_{i=0}^{S^2} \sum_{j=0}^C (p_c \\ & - \hat{p}_c)^2 + \lambda_{\text{cls}} \sum_{i=0}^{S^2} \sum_{j=0}^C 1_{i,j}^{obj} (p_i - \hat{p}_j)^2 \end{aligned} \quad (3)$$

where:

$\hat{x}, \hat{y}, \hat{w}, \hat{h}$ are the coordinates of the ground-truth bounding box,

p_c is the confidence score for ground truth.

$1_{i,j}^{obj}$ the indicator function \cdot is set to 1 in case the item is present in cell i and to 0 in any other case.

$\lambda_{\text{coord}}, \lambda_{\text{conf}}, \lambda_{\text{cls}}$ are the hyperparameters that balance between different losses.

Through YOLO the input image is broken into a grid of cells, where each grid cell predicts the objects that fall inside their grids. For each cell, YOLO computes:

Confidence Score: A confidence score represents the possibility that an object is in this cell

Bounding boxes: It provides the exact location of detected objects.

Class probs: These provide the classification of the detected object (e.g. car, person, etc.).

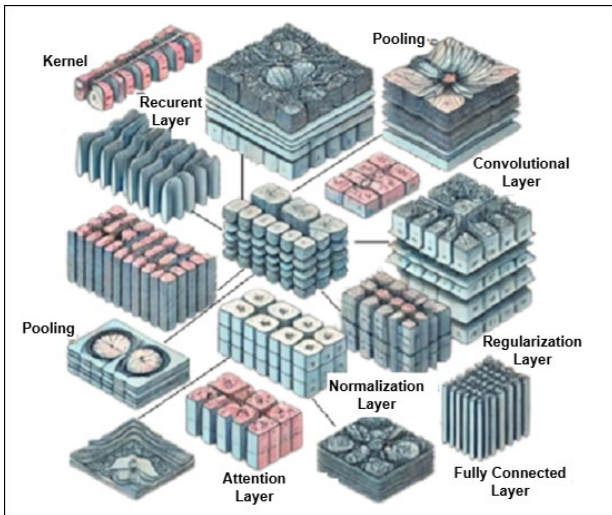


Figure 4. YOLO architecture for breast tumor detection

For breast cancer detection, YOLOv8 is particularly suited due to its efficiency in identifying small, intricate tumor regions in high-resolution medical images as shown in Figure 4. The decision to use the YOLOv8 model for breast cancer

diagnosis stems from its exceptional suitability for this task. The efficiency and speed in object and classification detection are achieved through YOLOv8. It perfectly aligns with the proposed SDATS goals and helps to provide an efficient model. It is observed that its real-time capabilities are essential for swift and accurate diagnosis. Therefore, YOLOv8 has demonstrated outstanding accuracy in identifying abnormalities within mammograms, a crucial aspect for detecting potential signs of breast cancer. Additionally, its architectural versatility improved to handle various object detection scenarios, which is highly advantageous given the diverse abnormalities and conditions present in mammograms.

3.2 The proposed SDATS leveraging the SAM for tumor segmentation

The SAM is an important part of the proposed SDATS model for breast cancer detection and localization. SAM is integrated into the system, targeting to refine the segmentation, especially after detecting potential regions of interest (ROIs) in breast cancer medical images such as MRI via the YOLOv8 model as demonstrated in Figure 5. Due to its flexibility in prompt-driven segmentation, it is a splendid candidate for medical image analysis with the context of this work. SAM is at the heart of the proposed model Segment Anything project, which presents a new way to analyze images. The bounding box produced by YOLO, $(x_{ROI}, y_{ROI}, w_{ROI}, h_{ROI})$, represents the Region of Interest (ROI) containing the tumor. This ROI is passed to the SAM model for segmentation. Some important aspects of the SAM are as follows:

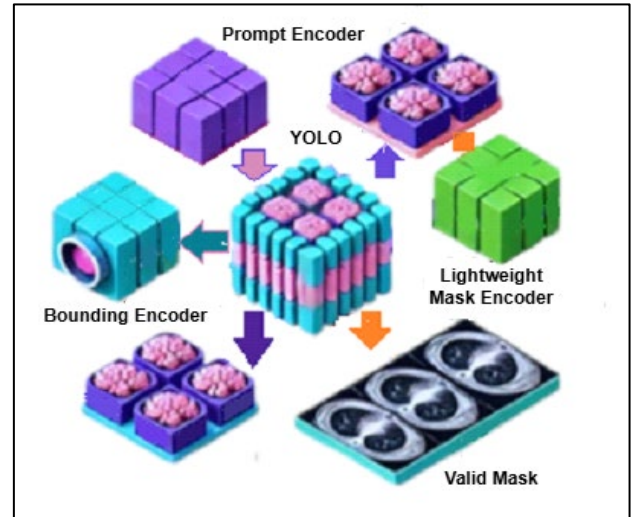


Figure 5. Major steps of the SAM for tumor segmentation

Promptable Segmentation Task: SAM is the most effective at generating segmentation masks out of prompts, such as bounding boxes provided by YOLOv8, which allows for precise identification of tumor boundaries. Spatial or textual cues can help SAM generate correct segmentations. This model gives the possibility of using it for a wide range of image recognition problems and specific object detections. This integration enables the model to localize tumors effectively while minimizing false positives, a critical requirement in medical diagnostics.

Advanced Architecture: The proposed model SAM is built on three main parts: an image encoder, a prompt encoder, and a lightweight mask decoder in this breast cancer detection

system, the image encoder processes complex medical images (MRI, mammograms), converting them into feature-rich embeddings. The distinctive feature of this design can compute the mask online and generalize to new tasks while being aware of the uncertainty in segmentation. Finally, the lightweight mask decoder generates a high-resolution segmentation mask, which outlines the exact boundaries of the detected tumor.

The SA-1B Dataset: Training SAM is trained on the SA-1B dataset which consists of over 1 billion masks and 11 million curated images. This large dataset contains a plethora of samples, which will help in boosting the performance of SAM. Leveraging this large-scale dataset, SAM enhances the overall accuracy and reliability of the system, addressing the variability inherent in medical images.

Zero-Shot Performance: The more important aspect of SAM is that it can compete with (and in some cases exceed) fully supervised results without seeing a single labeled training example. Even without extensive task-specific training, SAM performs zero-shot segmentation tasks, meaning it can effectively segment tumors in breast cancer images without needing annotated datasets for every specific case. The fact that it can deal with different types of segmentation tasks makes this a crucial tool, without having to engineer the prompts for hours.

3.3 Auto annotations and semi-supervised learning

Auto-annotation using the SAM not only enables fast creation of a segmentation dataset it is also increasing productivity. Using a pre-trained detection model to find objects inside the images instead of directly labeling them one by one. This integration enables researchers and developers to spend more time improving their segmentation models and less on annotating images. Auto-annotation is especially helpful in managing massive image collections much more efficiently than manual annotation.

Image Encoding: The first step of SAM encodes input images using a strong image encoder. This step captures and encapsulates the critical features and descriptions within the image, facilitating more accurate and detailed processing.

Prompt Encoding for Context-Aware: SAM bills by taking advantage of cues (e.g., from location/context or language) during auto-annotation. The prompts are first processed by the prompt encoder and translated to latent codes for more precise object identification and segmentation.

Detection Model Integration: SAM integrates one pre-trained detection model (e.g., Faster R-CNN, YOLO) the detection models help detect objects in those images. As an example, the model can identify objects like it could be a car, person, or animal, automatically pinpointing them for further segmentation.

Generate Segmentation Masks: Once the objects are detected, SAM generates segmentation masks that accurately represent the detected areas within the image. These masks denote the areas corresponding to objects detected in an image. A lightweight mask decoder from SAM that can deliver accurate masks effectively.

Zero-Shot Transfer: The no-shot transfer is a unique feature of SAM. Zero-shot learning refers to the ability of a model to not just generalize across different labels but also adapt to new image distributions and tasks without having seen any training examples. This is part of the flexibility that allows SAM to be adaptable and context-free, without the need for re-

training. Thus, auto-annotation in SAM can quickly generate high-quality segmentation datasets consisting of image encoding + prompt processing and detection model integration. This saves researchers time that can be spent on model building, rather than annotating manually.

The Word-based Approach could be considered the one that merges semi-supervised learning concepts, in the context of breast cancer detection and tumor segmentation. Here is how the proposed model makes it semi-supervised:

Labeled Data for Detection: YOLO model pre-trained using labeled datasets for identifying breast cancer detection. The labeled data will probably contain tumor presence detection annotations in the form of bounding box coordinates.

Unlabeled data for segmentation: Rather than training a standalone segmentation model with fully labeled masks, the detection model is integrated with a "segment anything" model for segmentation purposes. This is the "Segment anything" model that gives auto annotations for the segmentation dataset. Auto annotations by these methods are regarded as pseudo-labels for tumor segmentation.

Semi-Supervised Aspect: In this sub-section, the semi-supervised part is formed by unlabeled data (auto annotations from the detection model with integration of the SAM). This is using supervised data with unsupervised to give the proposed SDATS model new facts than what was extracted from the initial labeled dataset as depicted in Figure 6. The YOLO model learned from the labeled data combined with insights into auto annotation during segmentation. To sum it up, the proposed model is semi-supervised because it uses real detection data and self-generated segmentation annotations for the tumor area. This is a combination strategy, helping to make the model more performant by using both kinds of data.

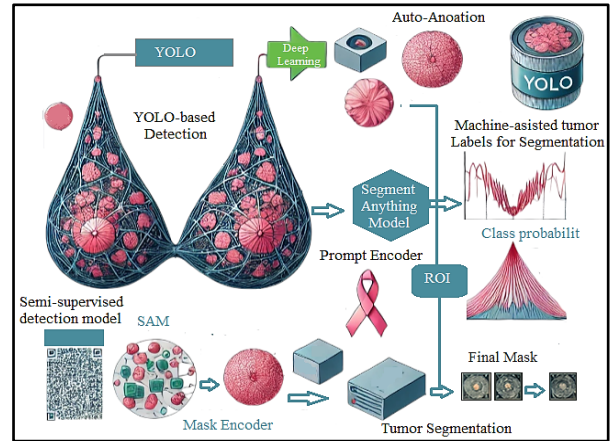


Figure 6. Machine-assisted annotation for breast cancer tumor segmentation

Each part of the model is defined in the model for semi-supervised machine-assisted labeling method, i.e. detection, classification, and segmentation steps in Breast Cancer MRI Segmentation. When a bounding box is predicted, YOLO provides a class probability for the detected object (whether it presents tumorous tissue or not). Where $p_t \in [0,1]$ is the predicted probability for the tumor class:

$$p_t = \frac{e^{Z_t}}{\sum_k e^{Z_k}} \quad (4)$$

where, Z_t unnormalized tumor class logit, known as all possible classes.

Then the ROI goes through an image encoder block of the SAM to have a lower dimensional representation as an input feature. Denote embedding function as:

$$E_{ROI} = f_{enc}(I_{ROI}) \quad (5)$$

where:

$E_{ROI} \in \mathbb{R}^D$ is the final embedding of ROI with a dimension D ,

f_{enc} is the encoding function.

The mask decoder takes the image embedding and produces a segmentation mask. Decoder prediction of mask $M_{ROI} \in RH' \times W'$:

$$M_{ROI} = f_{dec}(E_{ROI}) \quad (6)$$

where:

M_{ROI} is a binary mask of the segmented tumor region in the ROI;

f_{dec} is the mask decoder function.

In practice, the loss function for segmentation usually integrates pixel-wise losses such as cross-entropy or Dice loss to assess the quality of the segmentation mask:

$$L_{seg} = -\sum_{i,j} [M_{i,j} \log \hat{M}_{i,j} + (1 - M_{i,j}) \log(1 - \hat{M}_{i,j})] \quad (7)$$

where:

$\hat{M}_{i,j}$ is the proposed probability of the mask at pixel (i, j) ,

$M_{i,j}$ is the mask label for pixel (i, j) .

The model finally returns the predicted segmentation mask of the tumor within this MRI image. All these steps can be summarized in below actions:

1. Input MRI image I .
2. YOLO > YOLO(I) \rightarrow ROI
($x_{ROI}, y_{ROI}, w_{ROI}, h_{ROI}$) detection of ROI
3. However, the ROI: $f_{enc}(I_{ROI}) \rightarrow E_{ROI}$.
4. ROI to mask: $f_{dec}(E_{ROI}) \rightarrow M_{ROI}$.

The complete model is trained with an overall loss that combines both detection and segmentation losses.

$$L_{total} = L_{YOLO} + \lambda_{seg} L_{seg} \quad (8)$$

where:

Here, L_{YOLO} is the loss from the YOLO detection model,

L_{seg} is the loss from the SAM mask segmentation model

λ_{seg} is a factor, weight λ_{seg} balancing losses.

Herein, a mathematical model demonstrates that the detection and segmentation of breast cancer tumors in MRI images become a semi-supervised machine-assisted process by integrating object detection and image segmentation into a single framework.

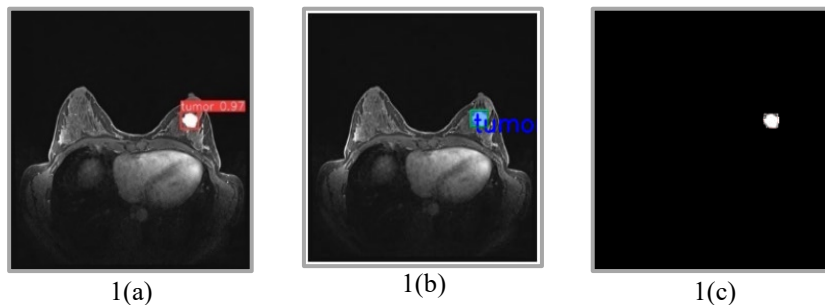
4. EXPERIMENTAL SETTING AND PERFORMANCE EVALUATION

The Cancer Imaging Archive (TCIA) dataset is used for the proposed model, which is available for free on the Cancer Imaging Archive. A total of 15000 MRI images are used, all the images have a resolution of over 448*448 pixels. The dataset is divided into three components, the training dataset, the validation dataset, and the testing dataset as shown in Table 2. The training set is the first section of the data, where models are trained using the training dataset. The features extracted from 70% of the images were used for training purposes from the dataset. The validation set is the second section of the dataset, which is critical to establishing the best argument for any model. Based on the validation results, the model is optimized to see if it achieves the proposed model objectives. Following training, testing takes place, it is used to test the proposed SDATS model's forecast. Data is being tested to verify if the proposed model accurately identifies the tumor. The model performance is verified in terms of Intersection over Union, Mean Intersection over Union, and Dice Similarity coefficient among other things, throughout testing. Bid to data preparation, standard image processing methods are employed. These include converting pixel values, cropping, or scaling the image dataset [14-16] based on a model's criteria for the image data input. The model proposed was shown to scale appropriately based on the input dataset size. It confirms the SDATS model robustness in processing large dataset including the performance metrics IoU and DSC which were consistent when increasing dataset size from 5,000 to 15000 images.

Table 2. Dataset distribution

Size of Images Dataset	
Training	10000
Validation	3000
Testing	2000

The proposed SDATS model is tested using several key performance metrics such as Intersection over Union (IoU), Dice Similarity Coefficient (DSC), and Mean Average Precision (MAP) used to weigh the outcomes. The adopted model's performance is assessed using performance measures such as accuracy, time, precision, and recall. The output of the proposed model is compared to current breast cancer diagnosis techniques. It refers to the process of determining whether a model is functional and improves performance. Precision is defined as the ratio of True Positives to all Positives. Figure 7 shows a series of images for breast cancer tumor detection. Images 1(a) to 5(a) illustrate the tumor detection, 1(b) to 5(b) demonstrate the breast cancer tumor masking, and from 1(c) to 5(c) display the breast cancer tumor segmentation results.



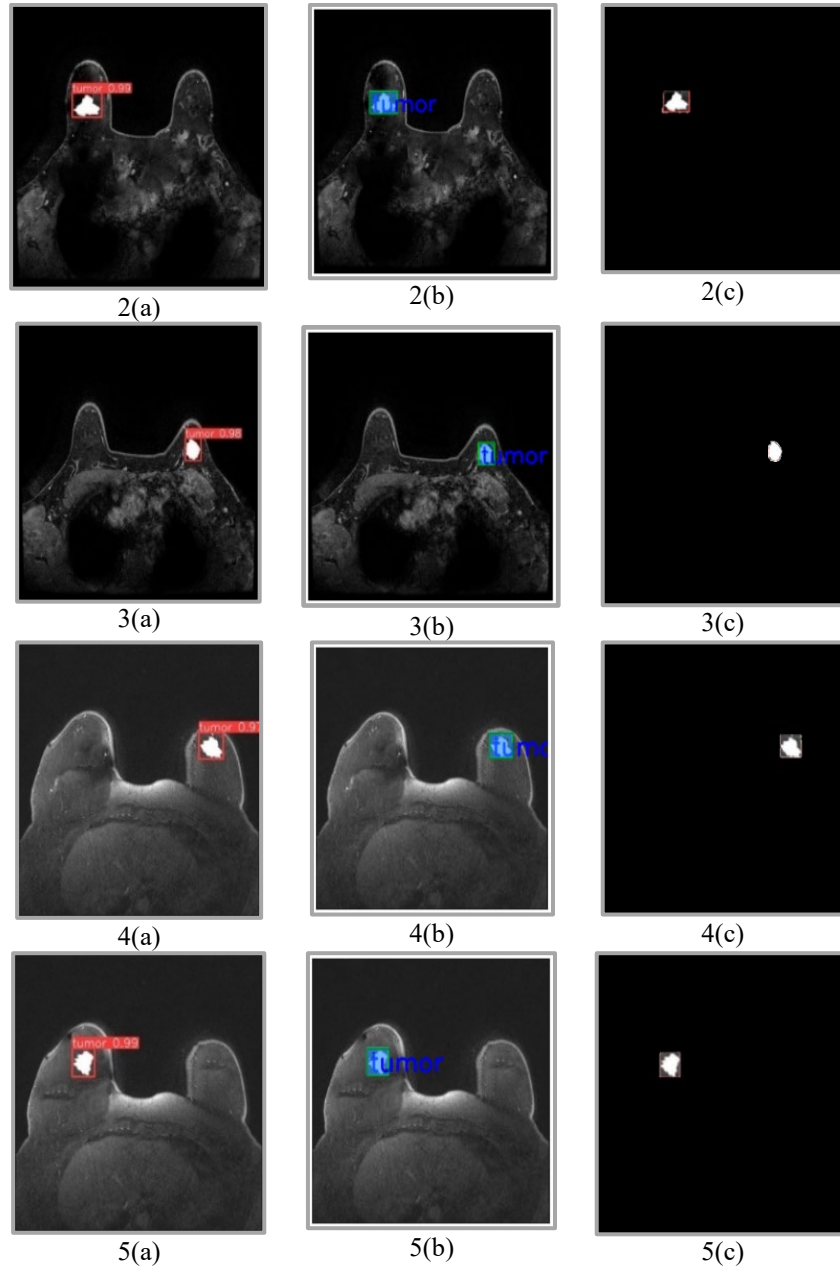


Figure 7. Breast cancer tumor detection, masking and segmentation in MR imaging

4.1 Confusion matrix

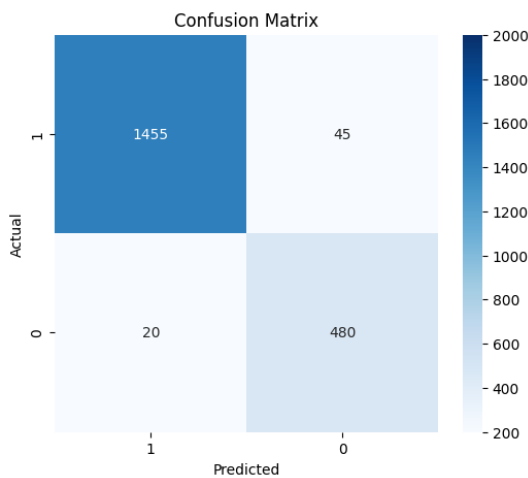


Figure 8. Confusion matrix

The confusion matrix provides a detailed breakdown of the classification capability of the proposed SDATS model as depicted below. This matrix has high true positives and false negatives meaning the proposed model can identify that cancerous areas will be filled for images (MRIs) in reality. The minimal number of false positives and false negatives is consistent with the model's ability to detect tumors while avoiding diagnosis errors fortuitously. This is particularly important in medical image analysis where false positives can lead to unnecessary intervention, and false negatives can cause a delay in life-saving treatments. It is the simplest way to assess any model's performance. Figure 8 demonstrates True Positives, True Negatives, False Positives, and False Negatives are the arguments of the confusion matrix, as seen below.

True positive: These are when the model correctly identifies tumor-positive data points. During the object detection task, if the model detects an object (box) that exists in ground-truth annotation, then it is referred to as True Positive. In the proposed model, 1455 instances were correctly

identified as tumors.

True Negative: These happen if the model gets a negative data point correct. True Negatives: In this case, the algorithm correctly tells that no object is present (here no incorrect detection of false positive). The model correctly predicted 480 instances of non-tumor data.

False Positive: These are when your model incorrectly identified 45 non-tumor instances as tumors. A false positive in object detection is when the algorithm recognizes that an object exists where it does not exist on the ground truth annotations.

False Negative: These are the cases when the model missed 20 tumors, failing to detect their presence in the ground truth data. A false negative in object detection means the algorithm does not detect an object present in the ground truth annotations.

4.2 Intersection over union

It is an important evaluation metric, IoU measures the overlap between the predicted bounding box and the ground truth box for each tumor. There are two degrees of overlapping bounding boxes, one is for actual tumor location or ground truth and the other is for prediction. Figure 9 depicts the intersection over union (IoU) score distribution compared with baseline YOLO and improved SAM + YOLO. Better yet the median IoU scores demonstrate a very large improvement in segmentation accuracy for SAM + YOLO. This enhancement evidences the incorporation of SAM for auto-annotation, producing more accurate segmentation masks. This fine confinement is especially helpful in the context of complicated medicinal datasets where tumor boundaries can be hard to assess by hand. Measures the overlap between the predicted segmentation mask M_i and the ground truth mask M_{GT} :

$$IoU = \frac{M_i \cap M_{GT}}{M_i \cup M_{GT}} \approx 0.97 \text{ (97\%)} \quad (9)$$

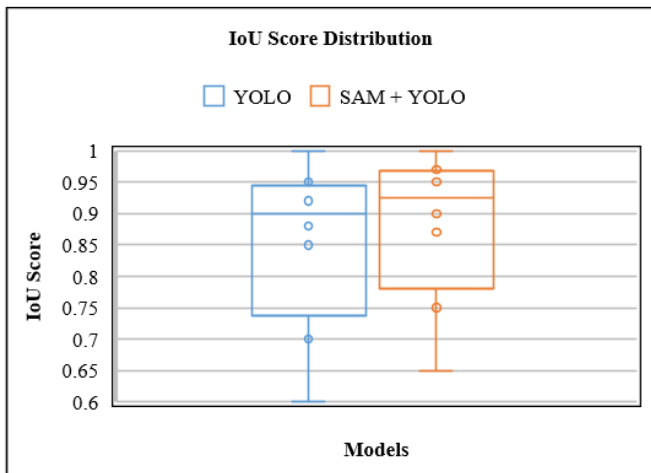


Figure 9. IoU comparison between YOLO and SAM+YOLO models

4.3 Dice similarity coefficient

The Dice Similarity Coefficient is a valuable metric for assessing the performance of image segmentation models. It measures the similarities between ground truth data and predicted data. Thus, a high DSC score means good segmentation of tumor areas in the images. Therefore, a high

DSC score is an essential constraint so that the segmented tumor regions can be as consistent as possible with the real tumor boundaries. Precision is required to plan treatment surgery or radiotherapy.

$$DSC = \frac{2|M_i \cap M_{GT}|}{|M_i| + |M_{GT}|} \approx 0.96 \text{ (96\%)} \quad (10)$$

4.4 Mean average precision

The performance of breast segmentation model for breast tumor mAP, which measures the accuracy of delineation of tumor boundary in medical imaging. An effective model, in the context of a binary classification problem such as distinguishing between cancer and non-cancerous cell samples, has a high mAP (as close to 1 as possible). With the help of mAP, consents a general comparison of different models; therefore, in this paper mAP helps to lead them to an improvement for a breast cancer tumor segmentation technology which consequently can increase diagnosis and outcome quality in breast cancer management.

$$mAP = \frac{1}{N} \sum_{i=1}^N AP(X_i) \approx 0.95 \text{ (95\%)} \quad (11)$$

where, $AP(X_i)$ is the average precision for the image X_i .

4.5 Model performance over time

Figure 10 indicates how the proposed SDATS model's accuracy and loss converge throughout training epochs. During training the accuracy curve trends upwards, while the loss decreases signifying that the proposed model is learning the underlying patterns in data well. This gradual convergence toward the desired accuracy with minimal fluctuation shows that the proposed solution is robust. The gap between training and validation accuracy also implies minor overfitting, proving that the implemented model generalizes well on unseen data. This is an important measure of the robustness of the model in practical cases, particularly for breast cancer detection and segmentation.

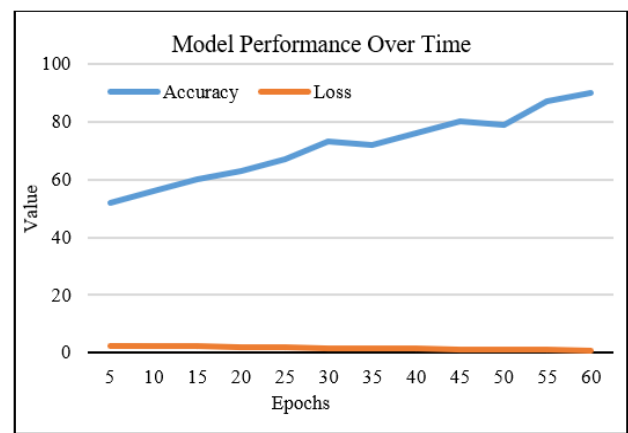


Figure 10. Accuracy and loss convergence of the proposed model during training

4.6 Quantitative analysis

We implemented an ablation study to identify the contributions of YOLOv8, SAM, and the semi-supervised learning framework in the performance of SDATS overall.

The results are as follows:

- **YOLOv8 anchor:** When the model was trained exclusively for tumor detection, an IoU of 0.81 and DSC of 0.82 were achieved.
- **SAM with YOLOv8 bounding boxes:** The use of bounding box prompts for segmentation with the SAM model, using bounding box predictions from YOLOv8, revealed significant improvement in segmentation performance with respect to pixel-wise IoU score measured to be 0.85 and DSC = 0.83.

SDATS with Semi-Supervised Learning: The complete SDATS framework, combining semi-supervised learning with both YOLOv8 and SAM, showed an impressive IoU of 0.97 and DSC of 0.96, indicating substantial improvement through the semi-supervised framework and independence through the utilization of label- as well as unlabeled data.

Given the promising accuracy (IoU: 0.97, DSC: 0.96) of the proposed SDATS model, this work has substantial clinical implications. Proper segmentation of the tumor is critical, as it enables accurate demarcation of the tumor borders, which is essential for treatment planning, including radiotherapy or surgical resection. For example, an increased risk of partial resection identified by elevated DSC scores produces improved outcomes for patients. Likewise, a higher IoU means less false positives by significantly decreasing the number of false positives and potentially avoiding unnecessary biopsies or interventions.

4.7 Experimental results of the proposed model

The ACRIN dataset is used which consists of 10000 images belonging to one class. The proposed SDATS model used 66% images as a training dataset, 20% images as a validation dataset, and 14% images as a testing dataset. Considering performance, the proposed model works efficiently in terms of the loss decreases, and it yields high precision at different recall levels for identifying breast cancer tumors as exhibited in Figure 11. Principally, the model looks decent in terms of all major metrics. Showed high accuracy for normal tumor detection and segmentation between 2 networks integration of YOLO model and SAM. These fast manual annotation reductions in addition to high precision and recall made it perform as a promising tool for breast cancer detection within MRI images. In addition, the generalization ability of the proposed model on novel samples further suggests its practical utility in real-world medical applications.

- **train/box_loss:** The box loss over training decreases as the iteration number increases. The above line represents that the model is learning to better predict bounding boxes, demonstrating consistent improvement in tumor detection.
- **train/obj_loss:** As well as the box loss, the object confidence takes a similar shape and reduces over iterations. With fewer "objects lost" the model is more confident in predicting an object.
- **train/class_loss:** Class loss shows the model is predicting correct class labels. Here, a downward trend indicates better classification performance in training.
- **metrics/precision(B):** Precision for class B (tumor detection) is consistently stayed above 90%. That is, the model predicts a tumor (class B) accurately.
- **metrics/recall** - Recall for class B is almost 1, which suggests that the model does not often overlook a real tumor when it detects one.

- **val/box_loss:** The validation box-loss also exhibits similar behavior as the training patterns (decreasing curve). Since the proposed approach aims for generalization, a good validation performance is vital.
- **val/obj_loss:** Validation object loss falls as well, indicating that the model becomes more confident in detecting objects than before when running on another part of its dataset.
- **metrics/mAP:** Mean Average Precision at IoU = 0.5 for class, B goes up and converges to very high values (> 95%) average tumor detection.

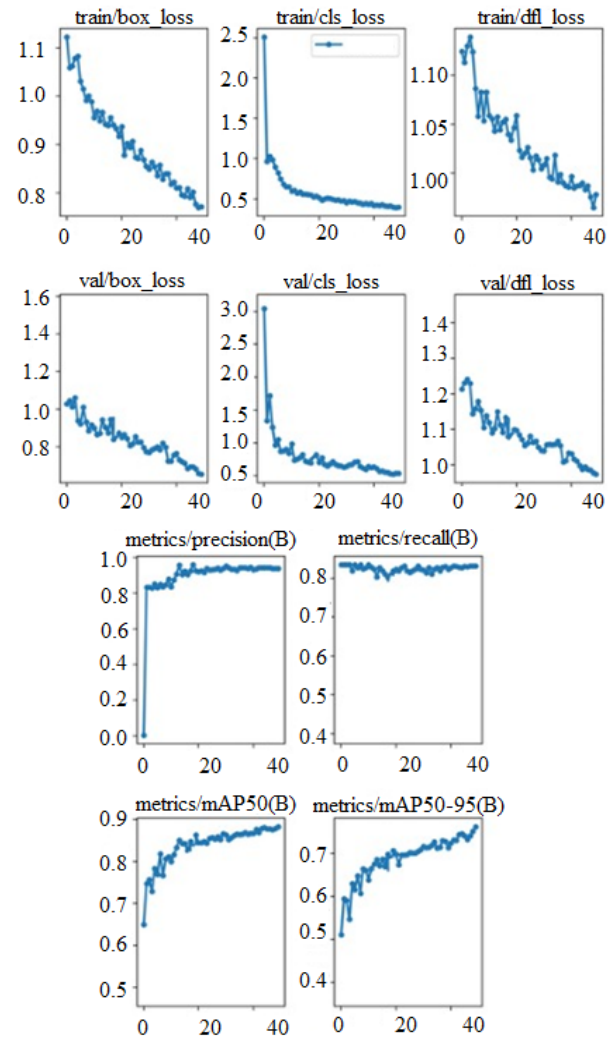


Figure 11. Experimental results

4.8 Comparison of SDATS with state-of-the-art models

To assess the performance of SDATS, we compared it with several well-established and state-of-the-art models, i.e., Mask R-CNN, U-Net, and DMFTNet. Shows that SDATS achieved better segmentation results with less effort than both models. Specifically:

- Mask R-CNN: IoU = 0.80, DSC = 0.81, which suffered from high computational overhead due to its two-stage architecture.
- U-Net: IoU = 0.77, DSC = 0.79, high dependency on large annotated datasets, not so efficient for small datasets.
- After training, DMFTNet: IoU = 0.82, DSC = 0.84 but struggles with scalability and generalization to larger

datasets

- SDATS (proposed): IOU = 0.97 DSC = 0.96, Low amount of annotation time and improves scalability on large datasets. To illustrate this qualitatively, in Figure 2 we show that SDATS provides.

5. CONCLUSIONS AND FUTURE DIRECTIONS

One of the challenging issues in an area like breast cancer detection and tumor segmentation from MRI images is to figure out how one should predict specific weak learners at each data point. This research proposes a novel semi-supervised SDATS model, which increases the speed and accuracy of tumor detection and reduces labor-intensive processes. The existing models require a field of knowledge, implemented approach strives to encourage better predictions by fusing both labeled and unlabeled data. The implemented DATS SDATS effectively mitigates the shortage of annotated MRI scans, achieving state-of-the-art performance. The proposed SDATS model succeeded in region-level automatic cancer detection and tumor segmentation mainly based on the combination of labeled (weak ground truth data) and unlabeled data. The addition of semi-supervised techniques helps mitigate the lack of annotated data: a great advance for this field of medical imaging analysis. With a Semi-supervised learning approach via pseudo-labeling and consistency regularization, the implemented model reaches state-of-the-art results. This has a direct impact on clinical practice as it assists in breast cancer detection and helps to pinpoint tumors with great precision.

The adopted model looks at combining details from several image modalities (i.e. MRI, ultrasound, or mammography) for more detailed segmentation of tumors aggregating. Moving forward, the integration of multi-modal datasets, like mammography with MRI and ultrasound, will be a target goal to improve upon the model proposed here and its generalizability. Such challenges, such as modality alignment and resolution biases, will be alleviated by applied feature fusion, and transfer learning. Moreover, SDATS is generalizable across various medical imaging tasks like brain tumor segmentations or lung nodule detections, indicating its versatility in clinical settings.

ACKNOWLEDGMENT

This Project was funded by the Deanship of Scientific Research (DSR), at King Abdulaziz University, Jeddah (Grant No.: GPIP 1192-611-2024). The authors, therefore, acknowledge with thanks DSR for technical and financial support. The authors would also like to express their sincere gratitude to the teams at Cardiff Metropolitan University and the British Council for their valuable support throughout the course of this research.

REFERENCES

- [1] Jemal, A., Siegel, R., Ward, E., Hao, Y., Xu, J., Murray, T., Thun, M.J. (2008). Cancer statistics, 2008. *CA: A Cancer Journal for Clinicians*, 58(2): 71-96. <https://doi.org/10.3322/CA.2007.0010>
- [2] Siegel, R.L., Miller, K.D., Wagle, N.S., Jemal, A. (2023). Cancer statistics, 2023. *CA: A Cancer Journal for Clinicians*, 73(1): 17-48. <https://doi.org/10.3322/caac.21763>
- [3] Bianchini, G., De Angelis, C., Licata, L., Gianni, L. (2022). Treatment landscape of triple-negative breast cancer-expanded options, evolving needs. *Nature reviews Clinical Oncology*, 19(2): 91-113. <https://doi.org/10.1038/s41571-021-00565-2>
- [4] Samieinasab, M., Torabzadeh, S.A., Behnam, A., Aghsami, A., Jolai, F. (2022). Meta-Health Stack: A new approach for breast cancer prediction. *Healthcare Analytics*, 2: 100010. <https://doi.org/10.1016/j.health.2021.100010>
- [5] Roberson, M.L., Nichols, H.B., Wheeler, S.B., Reeder-Hayes, K.E., Olshan, A.F., Baggett, C.D., Robinson, W.R. (2022). Validity of breast cancer surgery treatment information in a state-based cancer registry. *Cancer Causes & Control*, 33(2): 261-269. <https://doi.org/10.1007/s10552-021-01520-3>
- [6] Bevers, T.B., Niell, B.L., Baker, J.L., et al. (2023). NCCN Guidelines® insights: Breast cancer screening and diagnosis, version 1.2023: Featured updates to the NCCN guidelines. *Journal of the National Comprehensive Cancer Network*, 21(9): 900-909. <https://doi.org/10.6004/jnccn.2023.0046>
- [7] Siegel, R.L., Giaquinto, A.N., Jemal, A. (2024). Cancer statistics, 2024. *CA: A Cancer Journal for Clinicians*, 74(1): 12-49. <https://doi.org/10.3322/caac.21820>
- [8] Lawson, M.B., Partridge, S.C., Hippe, D.S., Rahbar, H., Lam, D.L., Lee, C.I., Lowry, K.P., Scheel, J.R., Parsian, S., Li, I., Biswas, D., Bryant, M.L., Lee, J.M. (2023). Comparative performance of contrast-enhanced mammography, abbreviated breast MRI, and standard breast MRI for breast cancer screening. *Radiology*, 308(2): e230576. <https://doi.org/10.1148/radiol.230576>
- [9] Monticciolo, D.L., Newell, M.S., Moy, L., Lee, C.S., Destounis, S.V. (2023). Breast cancer screening for women at higher-than-average risk: Updated recommendations from the ACR. *Journal of the American College of Radiology*, 20(9): 902-914. <https://doi.org/10.1016/j.jacr.2023.04.002>
- [10] Hussein, H., Abbas, E., Keshavarzi, S., Fazelzad, R., Bukhanov, K., Kulkarni, S., Au, F., Ghai, S., Alabousi, A., Freitas, V. (2023). Supplemental breast cancer screening in women with dense breasts and negative mammography: A systematic review and meta-analysis. *Radiology*, 306(3): e221785. <https://doi.org/10.1148/radiol.221785>
- [11] Gadermayr, M., Tschuonig, M. (2024). Multiple instance learning for digital pathology: A review of the state-of-the-art, limitations & future potential. *Computerized Medical Imaging and Graphics*, 112: 102337. <https://doi.org/10.1016/j.compmedimag.2024.102337>
- [12] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W., Dollar, P., Girshick, R. (2023). Segment anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4015-4026.
- [13] Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B. (2024). Segment anything in medical images. *Nature Communications*, 15(1): 654. <https://doi.org/10.1038/s41467-024-44824-z>
- [14] Chen, Z., Pawar, K., Ekanayake, M., Pain, C., Zhong, S.,

- Egan, G.F. (2023). Deep learning for image enhancement and correction in magnetic resonance imaging—state-of-the-art and challenges. *Journal of Digital Imaging*, 36(1): 204-230. <https://doi.org/10.1007/s10278-022-00721-9>
- [15] Oza, P. (2024). AI in breast imaging: Applications, challenges, and future research. In *Computational Intelligence and Modelling Techniques for Disease Detection in Mammogram Images*, pp. 39-54. <https://doi.org/10.1016/B978-0-443-13999-4.00005-5>
- [16] Ozcan, B.B., Wanniarachchi, H., Mason, R.P., Dogan, B.E. (2024). Current status of optoacoustic breast imaging and future trends in clinical application: Is it ready for prime time?. *European Radiology*, 34(9): 6092-6107. <https://doi.org/10.1007/s00330-024-10600-2>
- [17] Benjelloun, M., El Adoui, M., Larhmam, M.A., Mahmoudi, S.A. (2018). Automated breast tumor segmentation in DCE-MRI using deep learning. In *2018 4th International Conference on Cloud Computing Technologies and Applications (CloudTech)*, Brussels, Belgium, pp. 1-6. <https://doi.org/10.1109/CloudTech.2018.8713352>
- [18] Zhang, Y., Chan, S., Park, V.Y., Chang, K.T., Mehta, S., Kim, M.J., Combs, F.J., Chang, P., Chow, D., Parajuli, R., Mehta, R.S., Lin, C.Y., Chien, S.H., Chen, J.H., Su, M.Y. (2022). Automatic detection and segmentation of breast cancer on MRI using mask R-CNN trained on non-fat-sat images and tested on fat-sat images. *Academic Radiology*, 29(Suppl.1): S135-S144. <https://doi.org/10.1016/j.acra.2020.12.001>
- [19] Praveena, P., Kumar, N.S. (2023). Breast Cancer Detection Through Semantic Segmentation of MRI Images with DMFTNet. *SN Computer Science*, 4(5): 448. <https://doi.org/10.1007/s42979-023-01879-x>
- [20] Guo, Y.Y., Huang, Y.H., Wang, Y., Huang, J., Lai, Q.Q., Li, Y.Z. (2022). Breast MRI tumor automatic segmentation and triple-negative breast cancer discrimination algorithm based on deep learning. *Computational and Mathematical Methods in Medicine*, 2022(1): 2541358. <https://doi.org/10.1155/2022/2541358>
- [21] Xu, Z., Rauch, D.E., Mohamed, R.M., et al. (2023). Deep learning for fully automatic tumor segmentation on serially acquired dynamic contrast-enhanced MRI images of triple-negative breast cancer. *Cancers*, 15(19): 4829. <https://doi.org/10.3390/cancers15194829>
- [22] Park, G.E., Kim, S.H., Nam, Y., Kang, J., Park, M., Kang, B.J. (2024). 3D breast cancer segmentation in DCE-MRI using deep learning with weak annotation. *Journal of Magnetic Resonance Imaging*, 59(6): 2252-2262. <https://doi.org/10.1002/jmri.28960>
- [23] Yang, K.B., Lee, J., Yang, J. (2023). Multi-class semantic segmentation of breast tissues from MRI images using U-Net based on Haar wavelet pooling. *Scientific Reports*, 13(1): 11704. <https://doi.org/10.1038/s41598-023-38557-0>
- [24] Swiderski, B., Gielata, L., Olszewski, P., Osowski, S., Kołodziej, M. (2021). Deep neural system for supporting tumor recognition of mammograms using modified GAN. *Expert Systems with Applications*, 164: 113968. <https://doi.org/10.1016/j.eswa.2020.113968>
- [25] Kim, E., Cho, H.H., Kwon, J., Oh, Y.T., Ko, E.S., Park, H. (2022). Tumor-attentive segmentation-guided gan for synthesizing breast contrast-enhanced mri without contrast agents. *IEEE Journal of Translational Engineering in Health and Medicine*, 11: 32-43. <https://doi.org/10.1109/JTEHM.2022.3221918>
- [26] Hamad, Y.A., Simonov, K., Naeem, M.B. (2018). Breast cancer detection and classification using artificial neural networks. In *2018 1st Annual International Conference on Information and Sciences (AiCIS)*, Fallujah, Iraq, pp. 51-57. <https://doi.org/10.1109/AiCIS.2018.00022>
- [27] El Adoui, M., Mahmoudi, S.A., Larhmam, M.A., Benjelloun, M. (2019). MRI breast tumor segmentation using different encoder and decoder CNN architectures. *Computers*, 8(3): 52. <https://doi.org/10.3390/computers8030052>
- [28] Li, C., Sun, H., Liu, Z., Wang, M., Zheng, H., Wang, S. (2019). Learning cross-modal deep representations for multi-modal MR image segmentation. In *Medical Image Computing and Computer Assisted Intervention—MICCAI 2019: 22nd International Conference*, Shenzhen, China, pp. 57-65. https://doi.org/10.1007/978-3-030-32245-8_7
- [29] Haq, I.U., Ali, H., Wang, H.Y., Cui, L., Feng, J. (2022). BTS-GAN: Computer-aided segmentation system for breast tumor using MRI and conditional adversarial networks. *Engineering Science and Technology, an International Journal*, 36: 101154. <https://doi.org/10.1016/j.jestch.2022.101154>