Democratic and Popular Republic of Algeria
Ministry of Higher Education and Scientific Research

جامعة بجاية
Tasdawit n Bgayet
Université de Béjaïa

University A. Mira of Bejaia
Faculty of Exact Sciences
Department of Computer Science

MASTER'S THESIS

In
Computer Science

Option
Artificial Intelligence

Theme

TransUNet and adversarial learning for breast tumor segmentation in DCE-MRI

Presented by:
Aleaddine Bencheikh Lehocine
Salima Si Salem

Defended on June 29^{th}, 2025 before the jury composed of:

Chair	**Mr. Zair Bouzidi**	**Full Professor**	**U. A/Mira of Bejaia.**
Examinator	**Mr. Hachem Slimani**	**Full Professor**	**U. A/Mira of Bejaia.**
Examinator	**Mr. Foudil Mir**	**Associate Professor**	**U. A/Mira of Bejaia.**
Examinator	**Mr. Slimane Goudjil**	**Assistant Professor**	**U. A/Mira of Bejaia.**
Supervisor	**Mr. Fatah Bouchebbah**	**Associate Professor**	**U. A/Mira of Bejaia.**

Bejaia, 2024/2025.

⁂ *Acknowledgements* ⁂

First and foremost, all praise and thanks are due to Almighty
Allah,
for granting us the strength, patience, and perseverance to complete this work.

Our heartfelt thanks go to our parents, whose unconditional love, encouragement, and sacrifices have been the greatest source of motivation throughout our academic journey.

We would like to express our deepest gratitude to our supervisor, **Dr. Fatah Bouchebbah**, for his constant guidance, valuable advice, and generous support throughout the development of this thesis. His expertise and encouragement have been essential to our progress and success.

Finally, we are grateful to all those—teachers, friends, and colleagues—who contributed in any way to the completion of this project. Your support, whether big or small, is sincerely appreciated.

BENCHEIKH LEHOCINE AlaEddine , SI SALEM Salima

⁕ *Dédicace* ⁕

Je dédie ce modeste travail :

À mes parents bien-aimés, pour leur amour inconditionnel, leurs sacrifices silencieux, leur patience, leurs prières et leur soutien constant tout au long de mon parcours universitaire. Sans eux, rien de tout cela n'aurait été possible.

À notre encadrant ***Dr. Fatah Bouchebbah****, pour sa précieuse disponibilité, ses conseils éclairés et son accompagnement rigoureux tout au long de ce travail.*

À mes frères et sœur, pour leur encouragement moral et leur présence bienveillante.

À nos amis fidèles, pour leur soutien, leurs encouragements et les moments de complicité partagés qui ont rendu ce parcours plus agréable.

Et enfin, à tous ceux qui, de près ou de loin, ont contribué par un mot, un geste ou une pensée à l'aboutissement de ce projet.

Bencheikh Lehocine AlaEddine & Si Salem Salima

Contents

List of Figures

List of Tables

List of Abbreviations

AI: Artificial Intelligence.
BCE: Binary Cross-Entropy.
BD-RSIIM: Bi-directional Request-Supply Information Interaction Module.
BPE: Background Parenchymal Enhancement.
BRCA1: Breast Cancer Gene 1.
BRCA2: Breast Cancer Gene 2.
CAD: Computer-Aided Diagnosis.
CAM: Class Activation Mapping.
cGAN: Conditional Generative Adversarial Network.
CNN: Convolutional Neural Network.
CT: Computed Tomography.
DCE-MRI: Dynamic Contrast-Enhanced Magnetic Resonance Imaging.
DSC: Dice Similarity Coefficient.
EEB: Edge Encoder Block.
EGB: Edge Guidance Block.
ET: Echo Time.
FA: Flip Angle.
FCN: Fully Convolutional Network.
FID: Free Induction Decay.
FN: False Negative.
FP: False Positive.
GAN: Generative Adversarial Network.
HRT: Hormone Replacement Therapy.
IMIIN: Inter-Modality Information Interaction Network.
IoU: Intersection over Union.
MLP: Multilayer Perceptron.
MRI: Magnetic Resonance Imaging.
NMR: Nuclear Magnetic Resonance.
PD: Proton Density.

PET: Positron Emission Tomography.
PFC: Projet de Fin de Cycle.
PPV: Positive Predictive Value.
RF: Radio Frequency.
ROI: Region of Interest.
RT: Repetition Time.
SEN: Sensitivity.
SOTA: State-of-the-Art.
T1w: T1-weighted.
T2w: T2-weighted.
TN: True Negative.
TNM: Tumor, Node, Metastasis.
TP: True Positive.
TTA: Test Time Augmentation.
ViT: Vision Transformer.

General introduction

0.1 Scientific context

Breast cancer remains a leading cause of cancer-related mortality among women globally [21], making early and accurate detection a cornerstone of effective clinical management. In this context, advanced medical imaging modalities have become indispensable. Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI), in particular, has emerged as a highly sensitive tool for breast cancer screening, offering superior visualization of soft tissue and functional information about tumor vascularity that is often missed by other techniques.

Despite its sensitivity, the manual interpretation of DCE-MRI volumes is a demanding and time-consuming task for radiologists, prone to inter-observer variability and perceptual errors. This challenge has catalyzed the development of automated analysis tools, where image segmentation plays a pivotal role. By precisely delineating tumor boundaries, segmentation enables the quantitative analysis of crucial morphological characteristics, such as size, shape, and volume, which are vital for diagnosis, treatment planning, and monitoring therapeutic response.

The advent of deep learning has revolutionized the field of medical image segmentation. Architectures like Convolutional Neural Networks (CNNs), particularly the U-Net [164] model and its variants, have set new benchmarks by learning complex hierarchical features directly from image data. More recently, hybrid models like TransUNet [31], which integrate the global context-modeling capabilities of Vision Transformers with the local feature extraction strengths of CNNs, have demonstrated even greater potential for handling the intricate and varied nature of medical images.

0.2 Addressed problematic

Breast tumor segmentation in DCE-MRI remains challenging due to tumor heterogeneity, low contrast with surrounding tissue, and variability in imaging protocols. Traditional CNNs often fail to capture the long-range dependencies and contextual information necessary for precise tumor delineation. To address these limitations, this study proposes and empirically evaluates three methods: TransUNet, which leverages transformer-based self-attention for improved global context

modeling, and two cGAN-based frameworks that use adversarial learning to enhance segmentation accuracy and boundary refinement. By comparing these approaches, we aim to identify the most effective solution for robust and accurate breast tumor segmentation in clinical settings.

0.3 Aimed objectives and research methodology

To systematically investigate the aforementioned problematic, this project sets out to achieve the following primary objectives:

- **To implement and evaluate a strong baseline model** for breast tumor segmentation. This involves training a standalone TransUNet architecture in a purely supervised manner on the public BreastDM [175] dataset.
- **To develop and integrate two distinct cGAN frameworks** with the TransUNet generator. The first framework utilizes a standard PatchGAN discriminator, focusing on local, high-frequency details. The second employs a novel hybrid CNN-Transformer discriminator, designed to evaluate both local features and global contextual consistency.
- **To conduct a rigorous, empirical comparative study** of the three model configurations (Standalone TransUNet, TransUNet + PatchGAN, and TransUNet + Hybrid Discriminator). This ablation-style analysis aims to isolate and quantify the specific impact of adversarial training and discriminator design on segmentation performance.
- **To analyze the results both quantitatively and qualitatively.** Quantitative performance is measured using standard metrics (Dice and IoU). Qualitative analysis involves visual inspection of segmentation masks to understand model behavior, failure modes, and robustness to morphological variation.
- **To investigate the training dynamics** of the adversarial models by analyzing their loss curves, providing insights into the stability and effectiveness of the adversarial game.

0.4 Organization of the manuscript

This manuscript is structured into three main chapters, followed by a general conclusion.

- **Chapter 1** provides a general overview of the clinical context, introducing the fundamentals of breast cancer, the principles of biomedical imaging with a focus on MRI, and the foundational concepts of image segmentation.
- **Chapter 2** presents a state-of-the-art review of deep learning-based methods for DCE-MRI breast tumor segmentation. It discusses existing approaches, highlights their strengths and limitations, and identifies the key challenges that motivate this work.

- **Chapter 3** describes the second main contribution of this work. It details the methodology of our empirical study, including the dataset, the architectures of our three model configurations, the loss functions, and the training strategies. It then presents the complete experimental results, followed by a critical discussion of the findings and their implications.

- Finally, the **General conclusion** summarizes the key findings of this work, reiterates its main contributions, discusses the limitations of the study, and proposes promising directions for future research.

Chapter 1

General overview of breast cancer and image segmentation

1.1 Introduction

Breast cancer remains one of the leading causes of mortality among women worldwide. Despite extensive research, its exact causes are not fully understood. Hence, early detection and timely intervention remain the most effective strategies for improving survival rates. In this regard, biomedical imaging has become an essential tool in modern medicine, providing advanced technologies for detailed visualization of internal structures. These imaging modalities offer crucial insights for diagnosing, characterizing, and monitoring breast tumors.

Among the available screening and diagnostic techniques, Magnetic Resonance Imaging (MRI) is widely recognized for its superior sensitivity. Unlike conventional methods, MRI provides high-resolution imaging of soft tissues, making it particularly effective in detecting tumors that other modalities might miss. However, accurate interpretation of MRI scans relies on sophisticated image processing techniques, particularly image segmentation, which is vital for delineating tumor boundaries and distinguishing cancerous from healthy tissues. Specifically, image segmentation is considered as a key computational technique in medical imaging that enhances the accuracy of tumor identification and extraction. Therefore, it improves diagnostic precision and allows healthcare professionals to assess tumor size, shape, and progression more reliably.

Given these considerations, this chapter provides an overview of breast tumors, including their characteristics, diagnosis, and treatment. It also explores the fundamental principles of image segmentation, with a special focus on deep learning-based techniques. The chapter is structured as follows: Section 2.2 provides an overview of breast anatomy and the classification of breast tumors. Section 2.3 introduces biomedical imaging and examines key imaging modalities used for the detection of breast cancer. Section 1.4 outlines the most commonly adopted treatment approaches. Finally, Section 2.5 delves into the principles of image segmentation and highlights various seg-

mentation techniques applied in medical imaging. The chapter concludes with a summary of key findings in Section 2.6.

1.2 Anatomy and epidemiology

In this section, we present the anatomy of the breast and breast tumors. Additionally, we classify these tumors and conclude by discussing relevant statistical data at both the global level and in Algeria.

1.2.1 Normal anatomy of the breast

The breast is a paired organ located in the anterior and upper part of the thorax. It contains the mammary gland, which is responsible for milk secretion [168], and plays a crucial role in femininity. Functionally, it serves both nutritional and aesthetic purposes. Although the breast is predominantly developed in women, it also exists in vestigial form in men.

From an anatomical perspective, the female breast is composed of glandular, fibrous, and adipose tissues. It is positioned above the pectoralis major muscle and is anchored to the thoracic wall by Cooper's ligaments [155], which provide structural support. The soft consistency and shape of the female breast is mainly attributed to the layer of adipose tissue. The fibrous connective tissues and glandular structures contain lobules, which are responsible for milk production. Figure 1.1 illustrates the anatomical structure of the female breast.

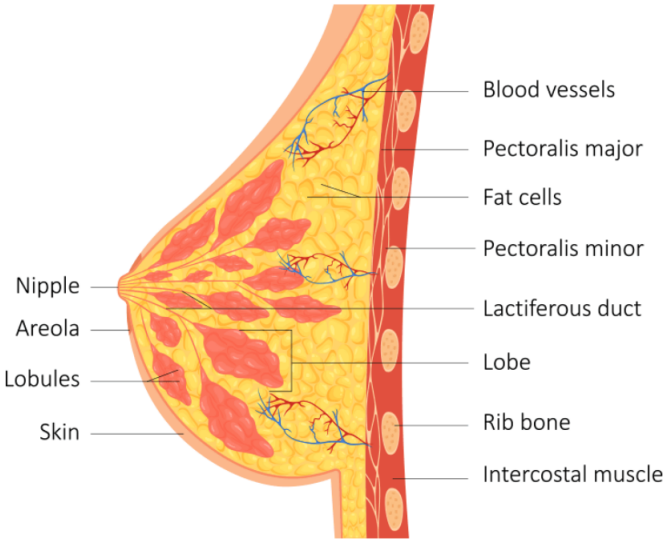

Figure 1.1: Illustration of a sagittal view of the anatomical structure of the female breast [189]

1.2.2 Breast tumor

A tumor is an abnormal growth of tissue caused by an abnormal cell proliferation. Tumors can be classified into two main categories: *benign* and *malignant*. Benign tumors are non-cancerous and generally do not spread to other parts of the body. In contrast, malignant tumors are characterized by uncontrolled growth and the potential to invade surrounding tissues or metastasize to distant organs, which defines them as cancerous.

On the other hand, breast tumors vary in nature depending on the type of tissue from which they originate. Some tumors result from benign breast conditions, such as fibrocystic changes, fibrosis, or localized lumps. These are usually not life-threatening and may not require aggressive treatment. However, some tumors are malignant (also called cancers) and typically develop in ductal cells (forming ductal carcinoma), lobular cells (leading to lobular carcinoma), or, in rarer cases, other breast structures. The majority of breast cancers originate in the milk ducts or lobules, as these areas are more susceptible to cellular abnormalities [80].

In addition, the onset of breast cancer is influenced by multiple factors, with genetic and hereditary predisposition playing a critical role. Individuals with a family history of breast cancer, particularly those with mutations in genes such as BRCA1 and BRCA2, are at significantly higher risk [130]. Beyond genetic factors, several environmental and lifestyle variables contribute to breast cancer incidence.

Also, key risk factors include hormonal, reproductive, and lifestyle influences. Early onset of menstruation (menarche) and late menopause increase a woman's lifetime exposure to estrogen, which has been linked to a higher risk of developing breast cancer. Similarly, reproductive factors such as having fewer children, late first pregnancy, and nulliparity (not having children) are associated with an increased risk. Exogenous hormone exposure, particularly through oral contraceptives and hormone replacement therapy (HRT), has also been implicated in some studies [23].

Furthermore, diet and body composition further contribute to risk. Research has shown that alcohol consumption and obesity, particularly weight gain in adulthood and increased visceral fat, can significantly elevate breast cancer risk. The role of nutrition, including the consumption of processed foods and high-fat diets, is also under investigation as a potential contributor [213].

On a positive note, certain protective factors have been identified. Breastfeeding is recognized for its protective effect, as it reduces lifetime exposure to estrogen and promotes cellular differentiation in breast tissue, thereby lowering cancer susceptibility. Additionally, regular physical activity has been linked to a lower risk of breast cancer by helping regulate hormone levels and maintaining a healthy body weight. These findings emphasize the importance of lifestyle choices in both cancer prevention and overall breast health [23].

Understanding these factors is crucial in developing effective breast cancer prevention strategies. While some risk factors, such as genetic predisposition, cannot be modified, lifestyle changes,

early screening, and advancements in medical research offer promising avenues for reducing breast cancer incidence and improving patient outcomes.

1.2.3 Classification of breast tumors

The classification of a tumor plays a crucial role in determining an accurate prognosis and developing an effective treatment plan. Among the various classification systems, the TNM (Tumor, Node, Metastasis) staging system is one of the most widely adopted for breast cancer. This system categorizes tumors based on three primary characteristics: the *size* of the tumor, the *extent of lymph node* involvement, and whether the *cancer has metastasized* to distant organs. Each of these factors contributes to defining the severity of the disease and guiding clinical decisions regarding treatment.

The size of the tumor (T) is a fundamental criterion in staging, as larger tumors are generally associated with a more aggressive disease progression. In addition to tumor size, the presence of cancer in lymph nodes (N) is another key factor. Lymph nodes serve as part of the body's immune defense, and when cancer spreads to these structures, the likelihood of further metastasis increases. The final criterion, metastasis (M), refers to whether cancer has spread beyond the breast and regional lymph nodes to other parts of the body, such as the liver, lungs, bones, or brain.

As the tumor progresses through these stages, the prognosis typically worsens. A small, localized tumor that has not spread beyond the breast has a much better prognosis than a tumor that has invaded multiple lymph nodes or metastasized to distant organs. The TNM classification system allows oncologists to determine the most appropriate therapeutic approach, which may include surgery, chemotherapy, radiation therapy, targeted therapy, or a combination of these treatments. By accurately staging the disease, clinicians can personalize treatment strategies and improve patient outcomes.

More details on the different stages of breast cancer tumors are illustrated in Figure 1.2.

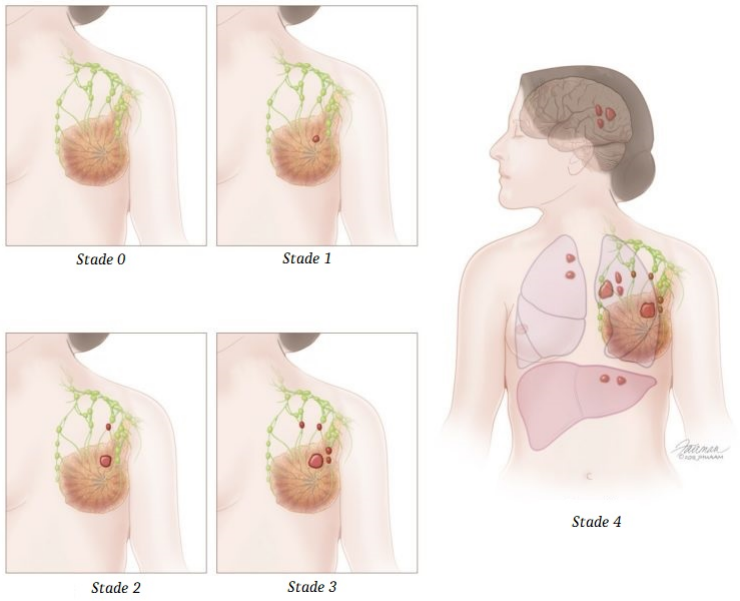

Figure 1.2: Illustration of breast tumor stages according to the TNM system [177]

1.2.4 Statistics on breast cancer

1.2.4.1 Worldwide statistics

According to Bray et al. [21], approximately 2.3 million cases of female breast cancer were diagnosed globally in 2022, making up nearly one in every four cancer cases among women. Consequently, breast cancer remains the most prevalent cancer among women and is the second most commonly diagnosed cancer overall. Figure 1.3 shows a pie chart representing the incidence rate of cancer worldwide in 2022, while Figure 1.4 illustrates the global mortality rate from cancer in the same year. Furthermore, the authors report that breast cancer is the most frequently diagnosed cancer in the majority of countries worldwide and continues to be the leading cause of cancer-related deaths in over 100 countries.

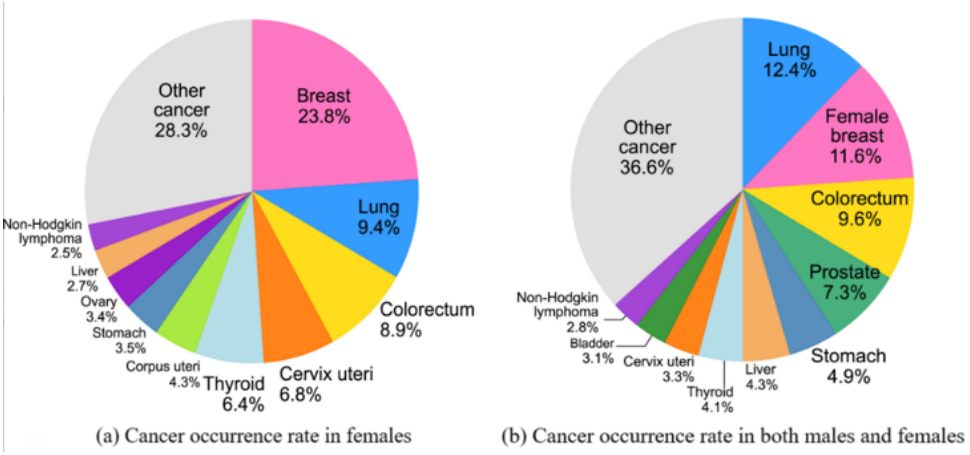

Figure 1.3: Pie chart of the global cancer incidence rate in 2022 [180]

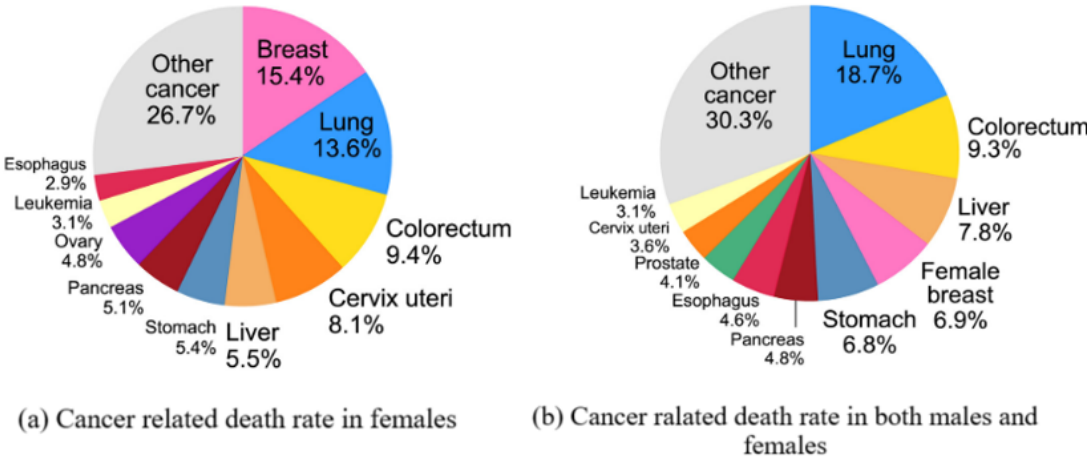

Figure 1.4: Pie chart of the global cancer incidence rate in 2022 [180]

Specifically, the incidence rate of breast cancer remains higher in developed countries such as Australia/New Zealand, European nations, and North America. In addition, over the past few decades, many transitioning countries in South America, Africa, and Asia have experienced a steady rise in the disease's incidence [21]. However, breast cancer mortality shows less variability compared to incidence rates, with the highest mortality observed in Melanesia, where Fiji continues to report among the highest mortality rates globally. Figure 1.5 illustrates a bar chart that shows age-standardized incidence and mortality rates by region for female breast cancer in 2022.

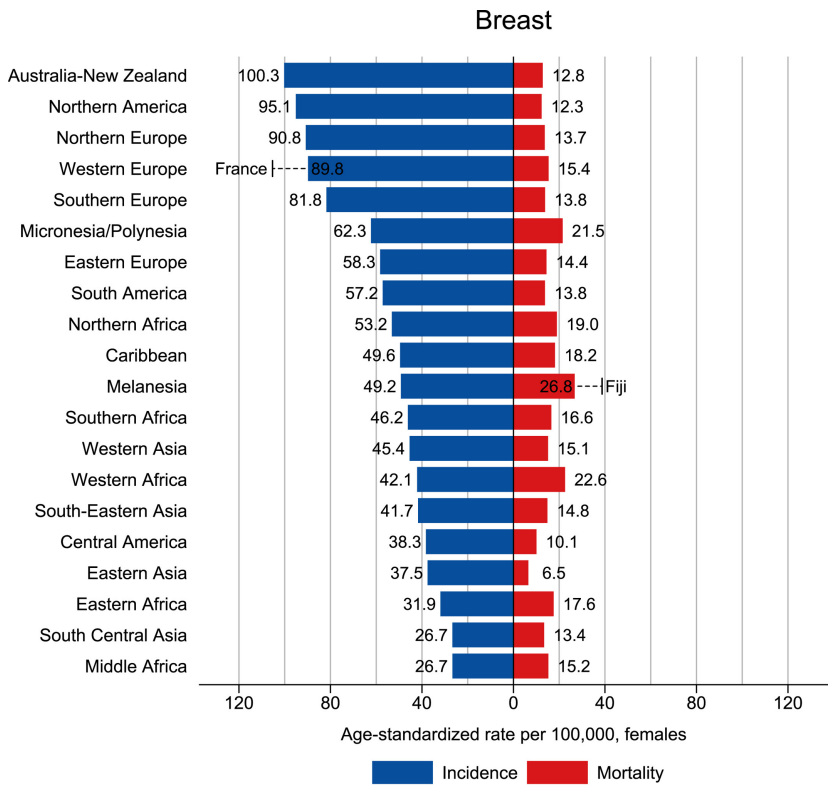

Figure 1.5: Bar chart of age-standardized incidence and mortality rates by region for female breast cancer in 2022 [180]

1.2.4.2 In Algeria

In Algeria, the latest statistics published in 2022 by the Global Cancer Observatory [65] of the World Health Organization [145] confirm that breast cancer remains the most frequently diagnosed malignancy, with 14,601 new cases reported in 2022, accounting for 22.6% of all cancer diagnoses. Furthermore, breast cancer constitutes a major public health concern as it represents the leading cause of cancer-related mortality among Algerian women, with 4,893 deaths recorded in the same year. Other prevalent cancer types include colorectal (7,747 cases), lung (5,040 cases), prostate (3,514 cases), and bladder (3,240 cases) cancers (see Table 1.1). These findings underscore the substantial burden of breast cancer in Algeria, positioning the country among the most affected in the African region.

1.3 Biomedical imaging for breast cancer screening

Biomedical imaging is a field of medicine that has witnessed significant advancements with the evolution of computing technologies. The terms "*biomedical imaging*" and "imaging diagnosis" refer to a generic, non-invasive process that allows for the observation of areas within the body that are not externally visible. This approach plays a crucial role in preventive health programs

Cancer	Incidence	Mortality
Breast	14 601	4893
Colorectum	7747	4380
Lung	5040	4599
Prostate	3514	1333
Bladder	3240	1818

Table 1.1: Incidence and mortality of the most prevalent cancers among men and women in Algeria (2022) [180]

and is widely regarded as the most effective method for the early detection of tumors. Moreover, the applicability and effectiveness of each imaging technique are heavily dependent on the specific characteristics of the organ being imaged. Therefore, selecting the most appropriate imaging technique prior to examining an organ is essential for accurate results [18].

The following subsections provide a brief overview of the primary imaging modalities used in radiology for breast imaging and the diagnosis of associated abnormalities, namely: *mammography*, ultrasonography, *Computed Tomography (CT)*, *Positron Emission Tomography (PET)*, and *Magnetic Resonance Imaging (MRI)*. These methods have become integral to the detection, diagnosis, and monitoring of breast cancer, offering valuable insights that aid in early diagnosis, treatment planning, and patient management. Each modality has its own strengths and limitations, and their combination often provides complementary information that enhances diagnostic accuracy.

1.3.1 Mammography

Mammography is a widely used mass screening modality that employs a non-invasive radiographic technique to capture the internal structures of the breast in an image [167]. Specifically, it involves exposing the breast to low-energy ionizing radiation (X-rays) at 30 kVp to produce images on planar radiographic films. Mammography is considered one of the gold standards in breast imaging due to its simplicity, speed, high specificity, and wide availability. Additionally, it has the ability to detect tumors that are not yet palpable, making it an effective tool for identifying abnormalities at the pre-clinical stage.

However, mammography has several limitations. Notably, the use of ionizing radiation over time could potentially contribute to the development of tumors [66]. Furthermore, the quality of the images produced is often inferior compared to other imaging modalities. Mammography also faces challenges in detecting abnormalities in women with dense breast tissue, where it may not provide accurate or sufficient results [88]. Despite these drawbacks, mammography remains a crucial tool for breast cancer screening, particularly when used in conjunction with other imaging techniques to enhance diagnostic accuracy.

1.3.2 Ultrasonography

Ultrasonography is also a non-invasive screening modality that uses sound waves to create a visualization of the breast. Specifically, it operates on the principles of echo emission and the transmission of ultrasonic waves (ranging from 2 to 20 MHz). Higher frequencies produce clearer images but penetrate less deeply into the patient. Ultrasound is particularly useful in diagnosing whether a nodule is solid or fluid-filled [12]. Additionally, it has the ability to detect tumors that may not be clearly identified through mammography, especially in cases involving dense breasts (rich in glandular tissue), such as in women under the age of forty.

However, due to the relatively low resolution of breast ultrasound images, false positive results may occur [105]. Therefore, ultrasound requires radiologists with proper training and specialized skills, including manual dexterity and keen observational ability [131]. Furthermore, the application of a coupling gel between the probe and the breast is necessary to eliminate air refraction effects, ensuring optimal image quality for accurate diagnosis. Despite these limitations, ultrasound remains a valuable complementary tool in breast cancer screening, especially when combined with mammography to improve diagnostic accuracy.

1.3.3 Computed Tomography (CT)

Similar to mammography, *Computed Tomography* (CT), also known as tomography by computer, uses ionizing radiation (X-rays). However, it allows for the production of tomographic slices (or planes) of the breast. As a result, CT provides a three-dimensional visualization of the organ [76]. Despite this capability, computed tomography offers limited advantages for breast cancer diagnosis, as it does not provide as much significant information compared to mammography. Furthermore, it complicates the acquisition process by increasing the dose of X-rays, which raises concerns about radiation exposure. Consequently, CT is rarely used as a primary modality for breast cancer screening but may be considered in certain clinical situations, such as when more detailed anatomical information is needed for staging or evaluating the extent of disease.

1.3.4 Positron Emission Tomography (PET)

Positron Emission Tomography (PET) is a nuclear medicine imaging modality in which a small amount of radioactive liquid material is injected into the body and used to diagnose various diseases, including breast tumors. Typically, in PET, a simple sugar is used as the radioactive substance, which is directly injected into the bloodstream. As a result, it accumulates in areas of the body where it releases energy in the form of gamma rays [36]. These gamma rays are detected by the PET scanner as signals, which are then converted by a computer into detailed images showing the functioning of tissues or organs.

More specifically, tumor cells require sugar for growth. Therefore, by monitoring how sugar is

utilized in the body (or sugar metabolism), the PET scanner is capable of detecting tumors. PET provides physiological information about the tumor being examined, in contrast to other imaging modalities that primarily capture morphological data. Furthermore, PET is often employed to diagnose or stage metastases or to search for contaminated lymph nodes [199]. This makes PET a valuable tool for assessing the metabolic activity of tumors and for detecting early-stage malignancies that may not be visible on other imaging modalities.

1.3.5 Magnetic Resonance Imaging (MRI)

According to Semchedine [**semchedine**], *Magnetic Resonance Imaging (MRI)* is a highly effective and entirely non-invasive imaging modality. In brief, the principle of this biomedical imaging technique is based on the magnetization phenomenon of hydrogen protons when exposed to an electromagnetic field. The MRI scanner primarily consists of a powerful magnet that emits electromagnetic waves. These waves induce the magnetization of hydrogen protons present in the tissues of the examined organ. During this magnetization state, the protons absorb energy and align with the direction of the applied field.

Once the electromagnetic wave emission ceases, the protons gradually return to their initial state, releasing the absorbed energy in the form of a radio signal, known as *Free Induction Decay* (FID). This signal is intercepted and measured by specialized receivers. The received signal is then processed by a computerized system to construct high-precision anatomical slices. These slices are subsequently assembled into a final MRI volume, providing detailed internal structural visualization of the targeted organ.

In 1971, Damadian [44] was the first to conceive the idea of a scanner using magnetic resonance to detect cancer in tissues. However, Lauterbur [101] was the first to utilize such a machine for diagnostic purposes, publishing the first clinical MRI image in 1973. A schematic representation of Damadian's machine is illustrated in Figure 1.6, while a depiction of a modern MRI scanner is shown in Figure 1.7.

1.3.5.1 MRI parameters and sequences

Various types of MRI sequences can be generated by modifying the sequencing parameters of the MRI system. To obtain the desired sequence, three essential parameters must be manually adjusted to their appropriate values using a dedicated control console. These parameters are as follows [169]:

- *Echo Time (ET)*: Refers to the time interval between the application of a radio frequency (RF) excitation pulse and the peak of the resulting FID signal.
- *Repetition Time (RT)*: Represents the time separating two successive RF excitations. It

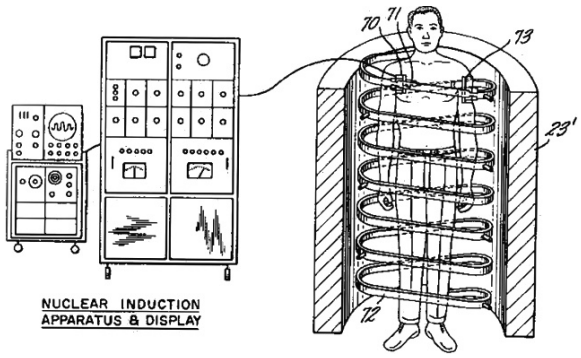

Figure 1.6: Schematic representation of Damadian's machine [43]

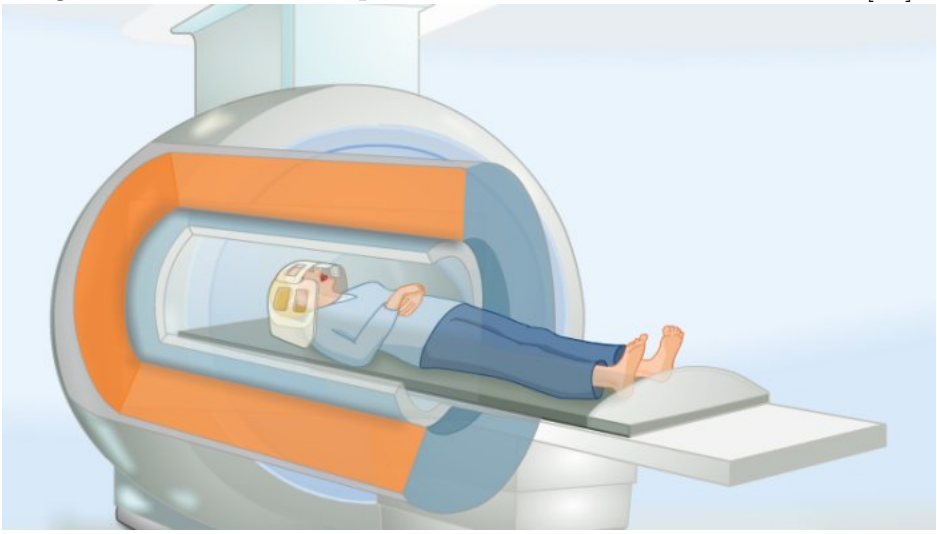

Figure 1.7: Representation of a modern MRI scanner [83]

determines the amount of magnetization recovered between each pulse. Both ET and TR are measured in milliseconds and are considered contrast factors.

- *Flip Angle (FA)*: This parameter controls the orientation of hydrogen protons in response to an RF excitation pulse.

Furthermore, the most commonly used MRI sequences for breast imaging are as follows [36, 169] :

- *Proton Density (PD)* : Obtained by using a short ET and a long RT.
- *T1-weighted (T1w)* : Achieved by using a long ET and a long RT. Useful for highlighting breast parenchyma or adipose tissue.
- *T2-weighted (T2w)* : Obtained using a short ET and a short RT. Effective for visualizing water-rich tissues such as cysts.

To obtain clearer images, a fat saturation technique is applied to suppress the signal from adipose tissue. This technique can be applied to both T1w and T2w sequences by utilizing the short relaxation times of adipose tissue [154]. When the fat signal is suppressed in a T1w image (or T2w), the resulting image is referred to as a fat-saturated T1-weighted image (or fat-saturated T2-weighted image).

1.4 Breast cancer treatment

After obtaining the results from medical imaging, the attending physician establishes a treatment plan based on the findings. The most commonly used treatments are surgery and therapy, but a combination of both can also be considered.

1.4.1 Lumpectomy

Lumpectomy is a breast-conserving surgical procedure in which the surgeon removes the tumor from the breast along with a surrounding margin of normal tissue (referred to as the "margin"). To perform this, the surgeon makes an incision ranging from two to seven centimeters on the breast and surgically removes the tumor along with a border of surrounding healthy breast tissue to ensure the complete removal of the affected area [186]. Post-lumpectomy, radiation therapy is required to eliminate any microscopic cancerous cells that may remain in the remaining breast tissue, thereby reducing the likelihood of local recurrence (reappearance of the same treated tumor).

1.4.2 Mastectomy

Non-conserving breast surgery, or *mastectomy*, involves the removal of all breast tissue from the chest wall [165]. Different types of mastectomy may be performed:

- *Total mastectomy*: The entire breast is removed, but the pectoral muscles beneath the breast tissue and the lymph nodes are preserved.
- *Modified radical mastectomy*: The entire breast is removed along with the affected axillary lymph nodes.
- *Skin-sparing mastectomy*: A total or modified radical mastectomy is performed, but most of the skin over the breast is preserved for reconstructive purposes.

It is important to note that mastectomy does not prevent cancer recurrence on the chest wall, and therefore, radiation therapy is recommended following the surgery [165]. Figure 1.8 illustrates the modified radical mastectomy.

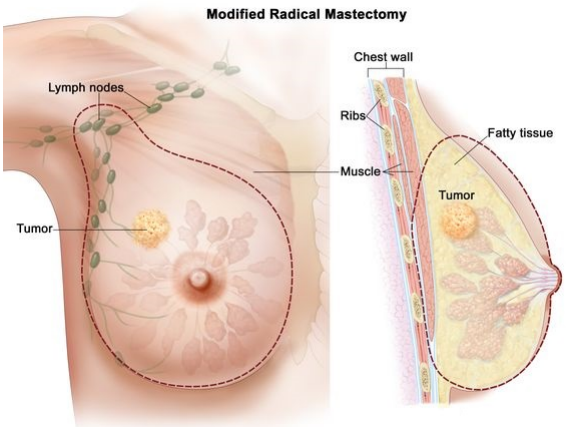

Figure 1.8: Illustration of the modified radical mastectomy [128]

1.4.3 Chemotherapy

Chemotherapy is a procedure that involves the use of anticancer drugs (or a combination of drugs) to eliminate cancer cells [165]. There are three main situations where chemotherapy may be used:

- *Before surgery (neoadjuvant therapy)*: In this case, chemotherapy is used either to slow the development of a tumor or to reduce the size of larger tumors before tumor resection.
- *After surgery (adjuvant treatment)*: Used to reduce the risk of cancer recurrence.
- *Metastasis*: To destroy cancer cells that may have spread to other organs in the body.

1.5 Image segmentation: principles and methods

In this section, we present the fundamental principles of image segmentation in image processing, along with an overview of a couple of well known deep learning based methods in the literature.

1.5.1 Image processing

Image processing is a fundamental discipline aimed at analyzing an image to extract meaningful information or provide a semantic interpretation of the objects it contains. It encompasses two primary types of processing: *low-level* and *high-level* processing.

Low-level processing algorithms focus on the relationships between the numerical values of image pixels without considering their real-world significance [90]. These operations typically include tasks such as noise reduction, contrast enhancement, and edge detection. In contrast, *high-level* processing algorithms analyze the information extracted by low-level techniques to symbolically

describe or interpret the content of the image [158]. This stage often involves object recognition, classification, and scene understanding.

The image processing workflow is generally structured into four key stages: *pre-processing*, *segmentation*, *feature analysis*, and *interpretation*. The first two stages (*i.e* pre-processing and segmentation) are considered low-level processing tasks, as they primarily deal with pixel-level transformations and region extraction. The last two stages (*i.e.* feature analysis and interpretation) fall under high-level processing, as they aim to derive meaningful insights from the segmented image, enabling decision-making in various applications such as medical imaging, remote sensing, and autonomous vision systems.

1.5.1.1 Definition of image segmentation

The term *image segmentation* refers to the process of partitioning an image into groups of pixels, also known as regions or segments, that share uniform characteristics. This process highlights each homogeneous region in a way that allows it to be easily distinguished from others [79]. The combination of the resulting regions reconstructs the entire image. Image segmentation is primarily used to detect Regions of Interest (ROIs) or to delineate the boundaries of regions within an image. For clarity, an illustration of an image and two possible corresponding segmented images are provided in Figure 1.11.

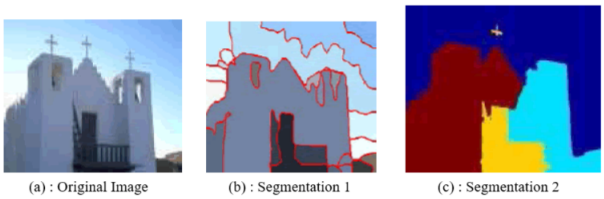

Figure 1.9: Illustration of an original image and its two possible segmented versions, demonstrating the image segmentation process [91]

1.5.1.2 Mathematical definition of image segmentation

A mathematical definition of image segmentation was proposed by Horowitz and Pavlidis [79]. Specifically, given an image Ω to be segmented, the segmentation process aims to partition the image into N sub-regions, denoted R_i, where $i = 1, 2, ..., N$ and $L(.)$ is a logical predicate that measures the homogeneity of a given region relative to a predefined set of features.

$$\Omega = \bigcup_{i=1}^{N} R_i, \tag{1.1}$$

$$R_i \cap R_j = \emptyset, \forall i, j \in \{1, 2, ..., N\}, \tag{1.2}$$

$$L(R_i) = \text{True}, \forall i \in \{1, 2, ..., N\}, \tag{1.3}$$

$$L(R_i \cup R_j) = \text{False}, \forall i, j \in \{1, 2, ..., N\}, i \neq j. \tag{1.4}$$

The definition includes four key conditions: Condition (1.1) implies that every pixel in the image is assigned to one region of the image, covering the entire image domain. Condition (1.2) states that each pixel must be assigned to exactly one region, ensuring mutual exclusivity between regions. Condition (1.3) requires that the resulting regions are homogeneous with respect to some predefined feature set (such as color, intensity, texture, etc.), meaning that pixels within a region share similar characteristics. Condition (1.4) indicates that the union of two distinct regions, each homogeneous by itself, results in a heterogeneous region. This condition implies that regions can be merged if and only if their combination results in a non-homogeneous region, reinforcing the segmentation as a process of finding meaningful boundaries in the image. These conditions form the foundation of a rigorous approach to image segmentation, ensuring that the segmented regions are both distinct and homogeneous.

1.5.1.3 Over-segmentation *vs* Under-Segmentation

The number of segments is a crucial criterion in assessing the quality of an image segmentation. An imbalance in the number of segments, either too many or too few, can lead to undesirable results. *Over-segmentation* occurs when the number of segments produced exceeds the expected number of segments [188]. This situation arises, for example, when a ROI in an image is divided into multiple semantically insignificant segments, instead of being segmented into a single connected region. In this case, the resulting segments are often too small or lack semantic coherence, making them inadequate for meaningful analysis. Despite being an undesirable outcome, over-segmentation is frequently employed in various research contexts as a preliminary step to simplify the computational complexity of more resource-intensive methods. Specifically, an over-segmented image, compared to its raw form, can be used as input for complex techniques, where each segment is treated as a superpixel. These adjacent and similar segments are then merged to produce a more accurate and fine-grained segmentation.

On the other hand, *under-segmentation* refers to a segmentation that produces fewer segments than expected [25]. This problem is generally more challenging to address than over-segmentation [28]. One common solution is to re-segment the under-segmented image; however, this approach may not always be suitable, particularly when the image contains overlapping segments. In such cases, the application of re-segmentation could exacerbate the issue, leading to over-segmentation. While both over-segmentation and under-segmentation are well-recognized phenomena in the

literature, there is no universally accepted method to resolve these issues. Nonetheless, over-segmentation is often preferred over under-segmentation because it serves as a useful preprocessing step before applying more refined segmentation techniques [16]. Over-segmentation generally provides a more detailed segmentation map, which can later be merged into more meaningful regions, improving the accuracy of the final segmentation.

1.5.1.4 Automatic and semi-automatic segmentation

Semi-automatic segmentation requires minimal user involvement in the segmentation process [20]. This type of segmentation is particularly useful when the user desires to control or guide the segmentation, such as specifying the ROI to be segmented or fine-tuning the precision of the obtained result [67]. The semi-automatic approach strikes a balance between automation and user control, allowing for customized intervention to improve accuracy or adjust parameters based on user expertise or specific needs.

In contrast, *automatic segmentation* aims to fully automate the process, utilizing advanced algorithms and prior knowledge to enable the computer to determine the final segmentation result without any human intervention. These methods rely heavily on machine learning, statistical models, or deep learning techniques to analyze image features and make segmentation decisions autonomously.

When comparing the two types of segmentation, semi-automatic methods are more widely presented in the literature, despite the growing interest in minimizing human involvement as much as possible in automatic segmentation methods [67]. The primary advantage of semi-automatic segmentation lies in its flexibility and adaptability, allowing users to intervene and refine results when necessary, while still benefiting from the efficiency of automation. However, automatic segmentation methods, driven by advancements in artificial intelligence and deep learning, are becoming increasingly capable of achieving high-quality results with minimal human oversight, making them an area of active research for a wide range of applications.

1.5.1.5 Objectives of image segmentation

Image segmentation is a low-level processing technique that serves as a fundamental and essential step in image analysis. Specifically, the goal of segmentation is to simplify the representation of an image into something more meaningful and easier to analyze [86]. By doing so, it reduces the complexity of subsequent high-level processing required to interpret the image. As a result, successful image segmentation often leads to more accurate and effective high-level results [160]. For instance, in pattern recognition, precise segmentation of a ROI enables the extraction of representative feature vectors such as size, shape, color, and location, which in turn facilitates the recognition process.

Due to its crucial role, image segmentation methods are widely applied across various fields,

including medical image processing, content-based image retrieval, remote sensing, and more. In medical imaging, for example, accurate segmentation of tissues and organs plays a pivotal role in diagnostic procedures, while in remote sensing, segmentation helps in identifying and classifying land cover types. The versatility and significance of segmentation make it a key area of research and application in computer vision.

1.5.2 Deep learning for image segmentation

Deep learning has revolutionized the field of medical image analysis, particularly in the segmentation of complex images. Traditional image processing techniques often fail to handle the variability and intricate details present in medical images. Deep learning models, by learning hierarchical representations directly from the data, provide more precise and automated segmentation. This advancement has a profound impact on clinical applications such as tumor detection, diagnosis, and treatment planning [103]. In what follows, we briefly describe four widely used deep learning methods for medical image segmentation in the literature.

1.5.2.1 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are a class of deep feedforward architectures that automatically learn hierarchical spatially localized feature representations directly from raw images [103]. By stacking convolutional layers with nonlinear activations and pooling operations, CNNs capture increasingly abstract patterns, ranging from simple edges to complex textures. Thus, making them highly effective for visual tasks, including medical image segmentation.

- **Convolutional layers:** These layers apply a set of learnable filters to the input image. Each filter performs a convolution operation, producing feature maps that capture essential local patterns such as edges, textures, and shapes. This operation is critical for extracting spatial hierarchies within the image [95].
- **Activation functions:** Non-linear activation functions (e.g. ReLU) are applied after convolutional operations to introduce non-linearity into the network. This step is essential for enabling the network to learn complex representations of the input data [103].
- **Pooling layers:** Pooling layers, such as max pooling, reduce the spatial dimensions of the feature maps. This reduction not only lowers the computational cost but also provides a degree of translation invariance, making the model more robust to variations in input images.
- **Fully connected layers:** In many classification tasks, fully connected layers are used to integrate features and make final predictions. However, in segmentation networks like U-Net, an encoder-decoder structure is employed to preserve spatial information, replacing the traditional fully connected layers [164].

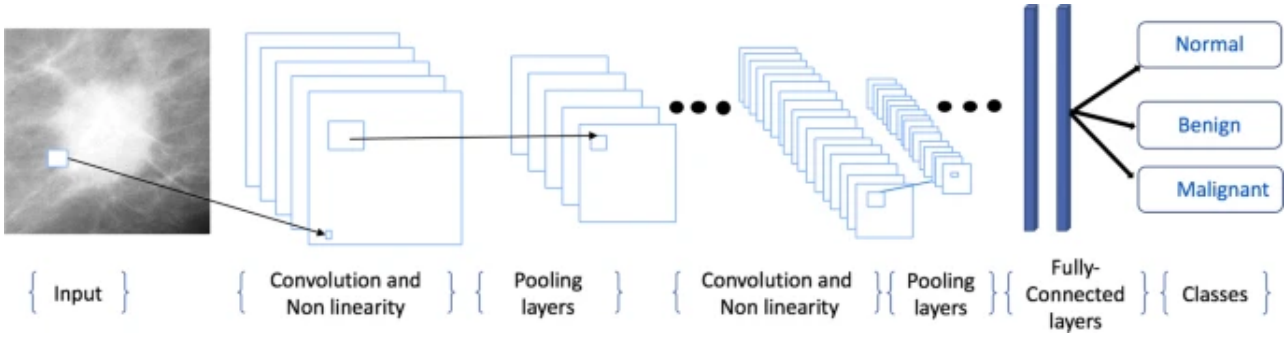

Figure 1.10: Example of a CNN architecture designed for tumor classification [1].

In medical imaging, segmentation is vital for accurately delineating ROIs, such as tumor boundaries in modalities like MRI, CT, and mammography. Accurate segmentation enables the extraction of critical information such as tumor size, shape, and location, which is essential for diagnosis and treatment planning. Advanced CNN architectures, including Fully Convolutional Networks (FCN) [119] and U-Net [164], have been specifically designed to meet the challenges of medical image segmentation. Comprehensive surveys have demonstrated that these deep learning models not only automate the segmentation process but also achieve performance levels comparable to human experts [113].

1.5.2.2 U-Net

U-Net is a fully convolutional neural network originally designed for biomedical image segmentation. Its strength lies in its ability to learn from very few training images while producing precise pixel-wise segmentation maps. The U-Net architecture consists of two main parts:

- **Encoder (contracting path):** This part is similar to a standard convolutional network. It applies a series of convolutional layers (typically with 3×3 kernels) followed by non-linear activations and max pooling operations. As the spatial dimensions decrease, the number of feature channels increases, enabling the network to capture the global context.
- **Decoder (expansive path):** The decoder upsamples the feature maps using up-convolution (or transposed convolution) to restore the original spatial resolution. At each upsampling step, features from the corresponding encoder layer are concatenated via skip connections, which help in recovering fine-grained spatial details.

The **skip connections** directly link corresponding layers in the encoder and decoder, fusing low-level spatial details with high-level semantic information. This U-shaped design allows the network to produce precise segmentation boundaries even when trained with limited data [164]. Figure 1.11 presents the architecture of the U-Net model.

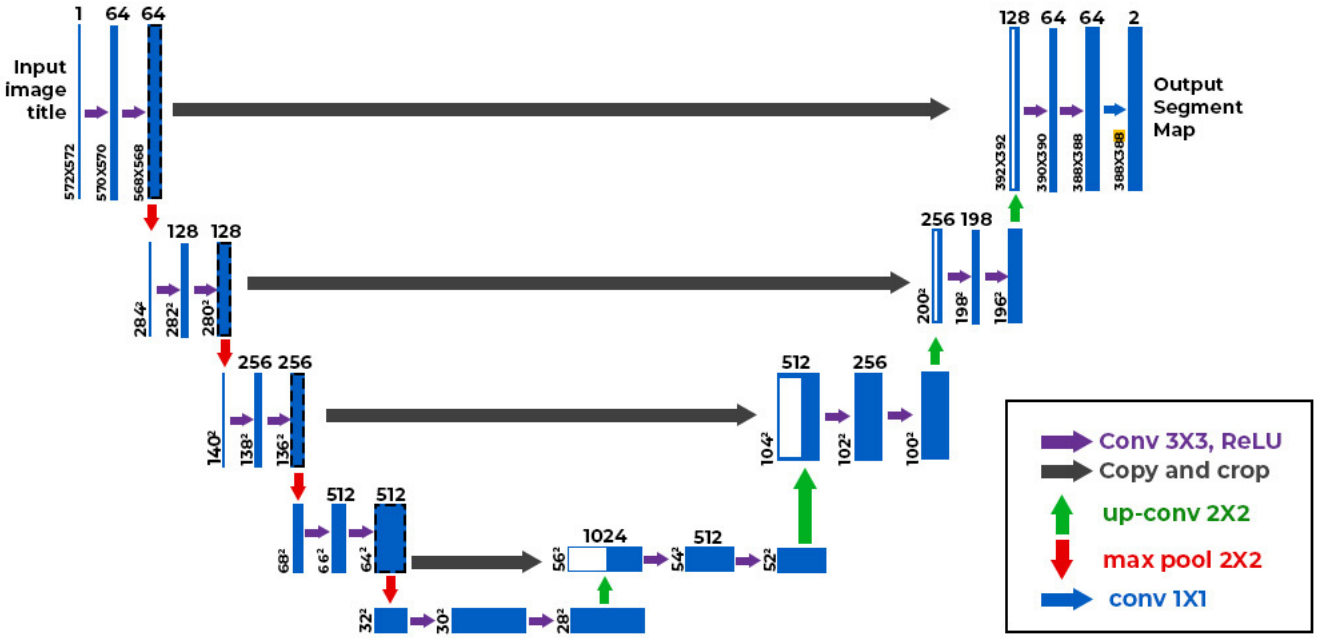

Figure 1.11: U-Net architecture (example for 32x32 pixels in the lowest resolution) [164]

1.5.2.3 Vision Transformer (ViT)

Vision Transformers (ViT) are a class of deep learning models that adapt the Transformer architecture, that is originally designed for natural language processing, to computer vision tasks, including image segmentation [48]. Unlike Convolutional Neural Networks (CNNs), which rely on local convolutional operations to extract spatial features, ViTs leverage self-attention mechanisms to capture global relationships across an entire image, offering a fundamentally different approach to processing visual data.

The ViT architecture begins by dividing an input image into a grid of fixed-size patches (e.g. 16x16 pixels). Each patch is flattened into a vector and linearly embedded into a higher-dimensional space, augmented with positional encodings to retain spatial information. These patch embeddings are then fed into a series of Transformer encoder layers. Each layer consists of two main components: a *multi-head self-attention* mechanism, which computes dependencies between all patches regardless of their spatial distance, and a *feed-forward neural network*, applied independently to each patch. Layer normalization and residual connections are used to stabilize training and improve gradient flow [185]. For segmentation tasks, ViT can be extended with a decoder, such as in models like SegFormer or TransUNet, to produce pixel-level predictions [192]. An example of the ViT's architecture is given in Figure 1.12 for visual illustration.

ViTs offer several advantages over CNNs, particularly for medical imaging applications like breast tumor segmentation. Their ability to model long-range dependencies makes them adept at capturing the global context of an image, such as the relationship between a tumor and surrounding

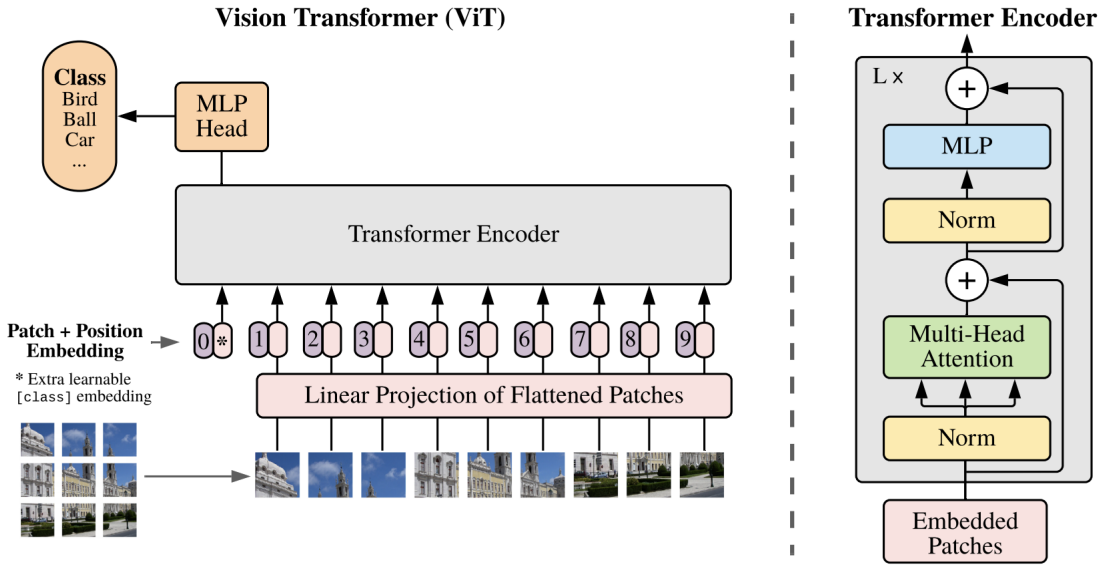

Figure 1.12: The Vision Transformer (ViT) architecture, where an input image is split into patches, embedded, and processed through Transformer encoder layers for classification or segmentation [48]

tissues in MRI scans, which may span beyond local convolutional receptive fields. Additionally, ViTs are highly flexible and can handle variable input resolutions when paired with appropriate positional encodings. However, they require large amounts of training data to outperform CNNs, as they lack the inductive biases (e.g. locality and translation invariance) inherent in convolutional operations. To mitigate this, pre-training on large datasets followed by fine-tuning on smaller medical datasets has become a common strategy [48].

In the context of breast cancer segmentation, ViTs have shown promise in recent studies. For instance, hybrid models combining ViTs with CNNs, such as TransUNet, leverage the strengths of both architectures: CNNs extract detailed local features, while Transformers model global interactions, improving tumor boundary delineation in complex MRI images [30]. While ViTs are computationally intensive due to their quadratic complexity with respect to the number of patches, advances like efficient attention mechanisms (e.g. Swin Transformer) are making them more practical for real-world applications [115]. This project could explore ViTs as an alternative or complement to U-Net, potentially enhancing segmentation accuracy by capturing both local and global tumor characteristics.

1.5.2.4 TransUNet

TransUNet is a hybrid deep learning architecture that integrates the strengths of Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) to enhance medical image segmentation, particularly for tasks like breast tumor delineation in MRI scans [32]. Originally proposed for 2D medical image segmentation, TransUNet combines the local feature extraction capabilities of U-Net with the global context modeling of Transformers, addressing the limitations of each approach when used independently.

The TransUNet architecture consists of three main components: an encoder, a bottleneck, and a decoder. The encoder employs a pre-trained CNN backbone (e.g. ResNet or EfficientNet) to extract low-level and mid-level features from the input image, capturing detailed spatial information such as edges and textures. These features are then transformed into a sequence of patch embeddings, similar to ViT, and processed through Transformer layers to capture long-range dependencies and global contextual information across the image. The bottleneck stage integrates these global features with local CNN features, enabling a richer representation of the image. Finally, the decoder, inspired by U-Net's architecture, upsamples the feature maps using convolutional layers and skip connections to reconstruct high-resolution segmentation masks, preserving fine-grained details like tumor boundaries [32] (see Figure 1.13 for illustration).

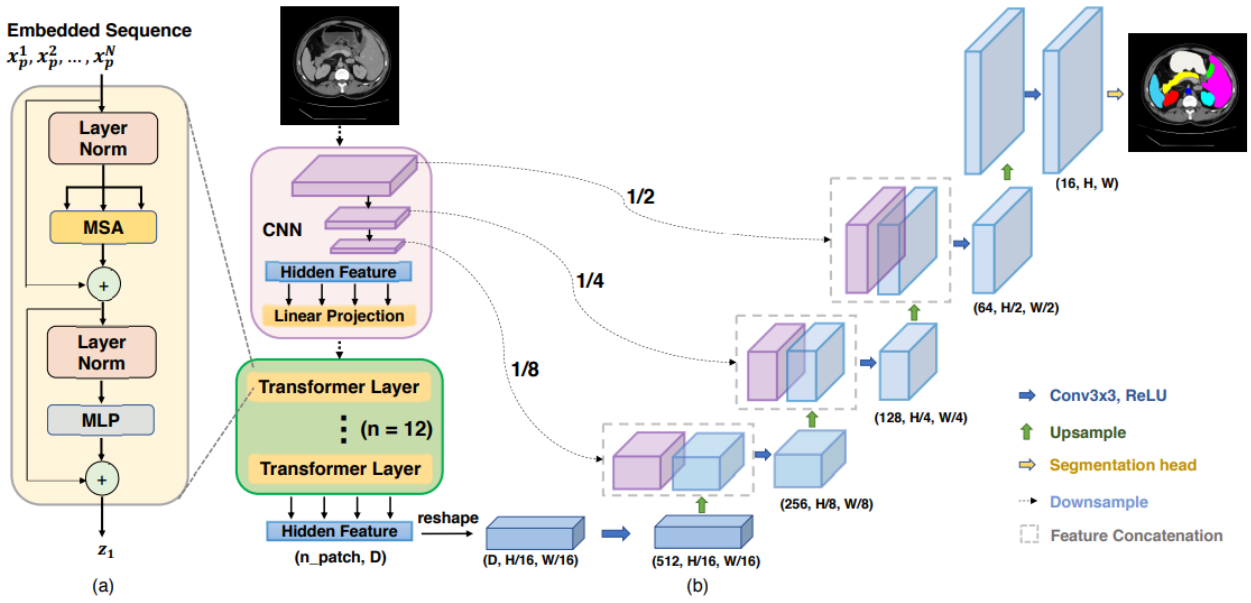

Figure 1.13: The TransUNet architecture, combining a CNN encoder, Transformer bottleneck, and U-Net-style decoder for medical image segmentation [30]

TransUNet offers significant advantages for breast tumor segmentation. Its CNN backbone excels at detecting local features, such as tumor edges and textures in MRI images, while the Trans-

former component models global relationships, such as the tumor's interaction with surrounding tissues or its overall shape. This hybrid approach mitigates the data inefficiency of pure ViTs, which require large datasets for training, and the limited receptive field of pure CNNs, which may miss global context. Studies have shown that TransUNet achieves state-of-the-art performance on medical segmentation benchmarks, including the Synapse dataset and breast cancer datasets, with improved Dice Similarity Coefficient (DSC) scores for tumor delineation [31].

In the context of this project, TransUNet is particularly promising for segmenting breast tumors in MRI scans. Its ability to combine local and global features could enhance the accuracy of tumor boundary detection, addressing challenges like irregular tumor shapes or low-contrast regions. However, TransUNet's computational complexity and memory requirements, due to Transformer layers, may necessitate efficient implementations or smaller models for real-time clinical applications. This project could explore TransUNet as a potential method, comparing its performance to U-Net or ViT on a breast cancer MRI dataset, potentially improving segmentation accuracy for clinical diagnosis and treatment planning.

1.6 Conclusion

The objective of this chapter was to provide an overview of breast cancer and image segmentation, serving as an introductory chapter to the broader topic. First, we defined breast cancer, biomedical imaging, and the existing treatment approaches used to combat this disease. We then introduced the fundamental principles of image segmentation and discussed some of the commonly used deep learning methods found in the literature.

In the next chapter, we will present a state-of-the-art review of various segmentation methods presented in the context of breast tumors DCE-MRI images, including 2D, pseudo 3D, and 3D approaches.

Chapter 2

State-of-the-art on DCE-MRI breast tumor segmentation

2.1 Introduction

Dynamic Contrast-Enhanced Magnetic Resonance Imaging (DCE-MRI) is a leading modality for breast cancer screening, providing detailed anatomical information alongside dynamic functional insight. However, its interpretation remains a complex and error-prone process when performed manually. Fortunately, image segmentation, a fundamental technique in computer vision, offers a promising solution by directing attention toward clinically relevant regions, such as tumors.Thereby enabling more efficient, accurate, and automated diagnosis.

In recent years, medical image segmentation has experienced significant advancements, largely driven by the rapid progress in deep learning technologies. As a result, various deep learning-based approaches have been proposed to automatically delineate breast tumor boundaries in DCE-MRI images. To offer a comprehensive understanding of these recent developments, this chapter provides a critical review and analysis of the most prominent state-of-the-art segmentation methods proposed within this scientific domain. The chapter is organized as follows: Section 2.2 presents the specifications related to the use of DCE-MRI in breast cancer screening and introduces the research problem addressed in this work. Section 2.3 provides a concise overview of recent deep learning-based approaches for breast tumor segmentation in DCE-MRI, as reported in the literature. Section 2.4 offers a comparative analysis of the reviewed methods, focusing on key aspects such as network architecture, dataset characteristics, and segmentation performance. Section 2.5 summarizes the findings, discusses current challenges, and outlines potential future research directions. Finally, Section 2.6 concludes the chapter.

2.2 Breast DCE-MRI and tumor segmentation

Dynamic Contrast-Enhanced MRI (DCE-MRI) for breast screening evolved from the broader development of MRI in medical imaging. The journey of DCE-MRI's development began with the discovery of Nuclear Magnetic Resonance (NMR) by Felix Bloch and Edward Purcell in 1946 [15]. Thereafter, Raymond Damadian proposed to use NMR for cancer detection in 1971, suggesting that malignant tissues exhibit distinct relaxation properties . This assumption led to the creation of the first clinical MRI scanner a couple of years later [102, 125]. For breast imaging, it is in the 1990s that MRI gained prominence as a complementary tool for high-risk screening. Particularly after the introduction of DCE-MRI. Specifically, the latter involves injecting a gadolinium-based contrast agent to enhanced the visibility of regions of interest in the resulted DCE-MRI images, such as tumors.

Nowadays, DCE-MRI is extensively used to screen the breasts for tumor analysis for its remarkable high sensitivity, especially when the patient presents dense breasts or a high-risk of tumor. Specifically, DCE-MRI provides exhaustively detailed view of tumor morphology, vascularity, and kinetic patterns. In addition to this, it provides 3D high-resolution visualizations of the both breasts. Such important features, and several others, are not present in mammography or ultrasonography which are conventionally used in clinical routines. [97]. Figure 2.1 presents an illustration of the MRI scanner used to scan the breasts.

2.2.1 Analysis of DCE-MRI images to diagnose breast tumors

The *double reading* procedure is the methodology used in the traditional routine to analyze DCE-MRI breast images [7]. Explicitly, two radiologists are charged to individually visualize the medical image to analyze and to write his/her interpretations of the findings in a report. If there are disagreements on interpretation, the two involved radiologists can either have a discussion about their opinions or seek advice from a third expert. However, like any human being, radiologists can face errors due to factors such as exhaustion, inattention, or even imperfections in the analyzed image due for instance to patient movement during the DCE-MRI exam. Consequently, the double reading procedure is not fully immunized. In this context, Nodine and Kundel [143] describe three main errors that can occur:

- A *visual error* that occurs when the presence of a tumor is not noticed. Usually because of a selective reading to focus on parts of the image that thought to be more interesting to explore.
- A *recognition error* that happens when a tumor is actually noticed, yet misidentified.
- A *decision-making error* that occurs when a tumor is identified, but misclassified.

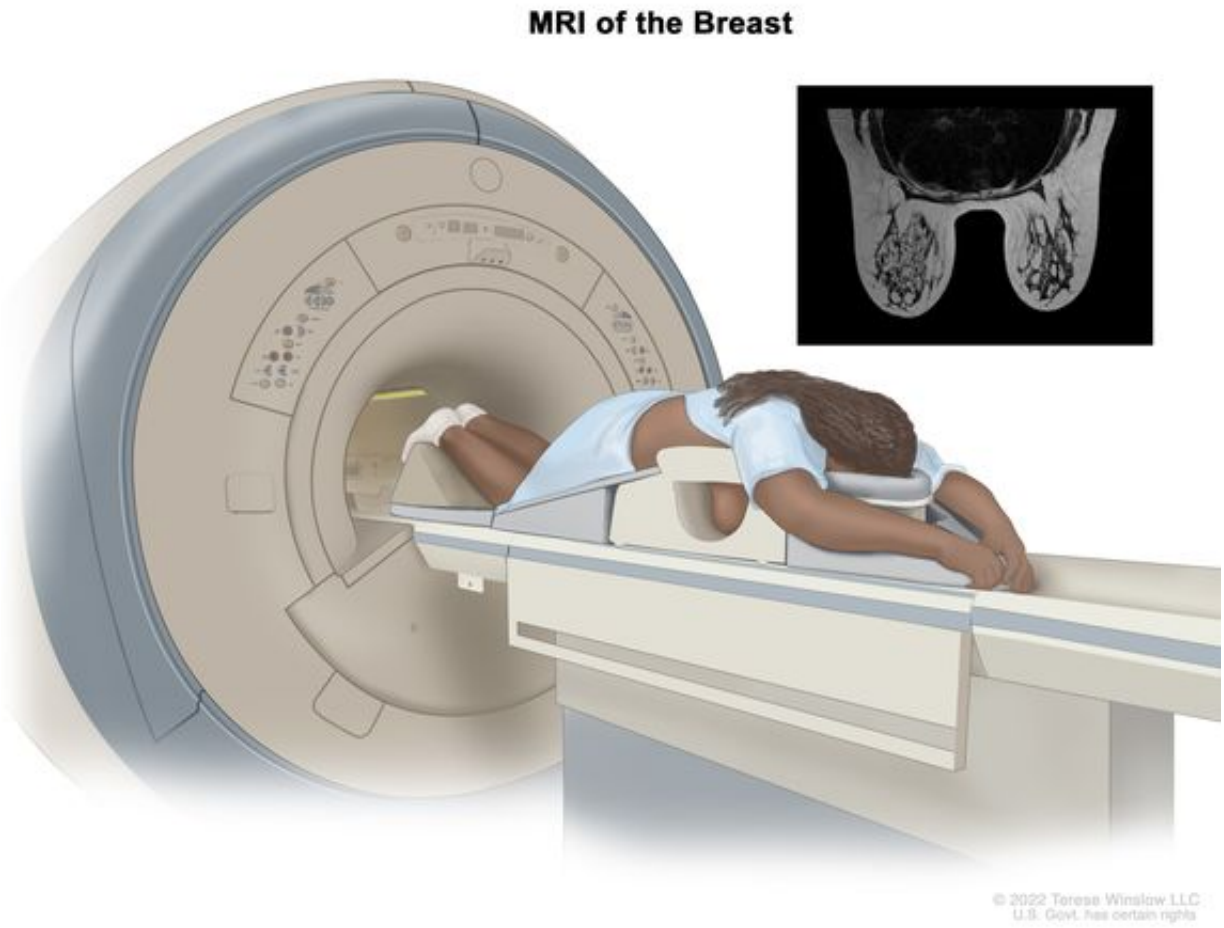

Figure 2.1: Illustration of a breast DCE-MRI scanner for cancer diagnosis [204]

Besides the possibility of error, DCE-MRI also imposes another challenge. Actually, the search area is large because the modality creates 3D scans with hundreds slices to explore in the axial, sagittal, and coronal planes. Making the exploration process totaly exhausting and time consuming.

On the other hand, the integration of computer science, and more precisely artificial intelligence, in modern radiology has had a significant impact on dealing with these hurdles [135]. Specifically, artificial intelligence proposes a variety of computer methods that aid in the analysis of medical images to establish more reliable diagnoses, such as *Computer-Aided Diagnosis* (CAD) system. Specifically, a CAD system is a computer software that is specially conceived to help determine the type, severity, stage, progression, or regression of a tumor [37]. The *US Food and Drug Administration* approved the use of this system in 1998, since then it has been widely used as a second reader to interpret medical images [118]. Moreover, the advantages of such a system in diagnosing breast tumors in DCE-MRI, especially in dense breasts, have been demonstrated in several studies [62, 118, 201, 107]. Technically, a CAD system is made up of four main steps that

are carried out sequentially, namely: the *preprocessing* step, the *segmentation* step, the *features extraction and selection* step, and the *classification* step.

2.2.2 Importance of DCE-MRI breast tumor segmentation

Segmentation, as described in the previous chapter, is the process of distinguishing regions of particular interest from other regions in an image. Hence, in the case where the delimited regions are tumors, segmentation allows not only to delicately define the boundaries of the contaminated areas, but also offers a possibility to inspect their appearances [141]. Therefore, segmentation represents a critical step in a breast DCE-MRI CAD system since its result significantly affects the performance of the next steps. Specifically, the more accurate the tumor segmentation, the more discriminatory features can be extracted, hence the more realistic the prognosis can be [78].

Moreover, DCE-MRI breast tumor segmentation used as a standalone process, outside the CAD system framework, offers several benefits targeting two important healthcare aspects, namely: enhancing patient outcomes and maintaining breast esthetics. For patient outcomes improvement, segmentation identifies contaminated zones at their earliest stages (when they are likely not visible visually), and/or maps their size, shape, and location. As a result, it decreases unnecessary biopsies and prevents the development of aggressive breast cancers by enabling oncologists to customize timely treatment plans and to monitor the tumor. Particularly when there is a suspected return after ablation or to follow up changes in growth or response to treatment.

On the other hand, the esthetic aspect of DCE-MRI breast tumor segmentation resides is how it can improve surgical precision while maintaining the natural appearance of the breast. Explicitly, the use of segmentation to accurately identify tumor boundaries allows to remove only affected tissue of the breast. Which helps surgical teams avoid unnecessary excision to reduce breast asymmetry and hence the need for extensive reconstruction. In addition, it enables less invasive procedures, resulting in smaller scars and improved cosmetic results.

2.2.3 Challenge of breast DCE-MRI images segmentation for accurate extraction of tumors

Considerable efforts have been made to present the most effective method for optimal DCE-MRI breast tumor segmentation so far. Yet, even with many proposed approaches and registered appreciable performance scores in the literature, there is no clearly described final solution. Mostly, this is due to many facts and challenges that are acting as roadblocks in the development of a final algorithm. Bellow, we describe this challenges according to the different cases: *benign tumors*, *malignant tumors*, and *normal (healthy) breast tissue*.

- **Benign tumors:** Such as fibroadenomas or cysts. They typically appear as well-defined, smooth-edged masses on DCE-MRI. They often show slow, gradual, or no significant en-

hancement in the early phases, with minimal or no washout in late phases, reflecting low vascularity and stable perfusion [81]. These tumors pose segmentation challenges due to their similarity to normal tissue in contrast, requiring algorithms to distinguish subtle enhancement patterns without over-segmenting healthy areas (*i.e.* false positives). Figure 2.2 presents an example of a benign tumor of BreastDM dataset [211] showcasing an overlapping contrast with its surrounding uncontaminated tissue of the breast.

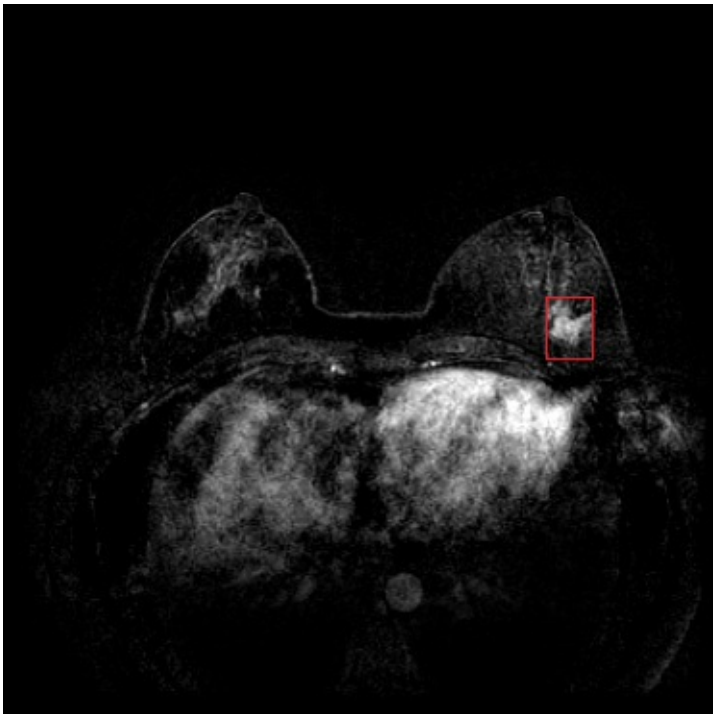

Figure 2.2: An illustration of a benign breast tumor highlighted with a bounded box in DCE-MRI taken from BreastDM dataset [211]

- **Malignant tumors:** Such as invasive ductal carcinoma or invasive lobular carcinoma. They exhibit rapid and irregular enhancement in the early DCE-MRI phases due to increased angiogenesis and vascular permeability, followed by a rapid washout in later phases [98]. These tumors often have irregular, spiculated margins and heterogeneous enhancement. Making them distinct from benign lesions but challenging to segment accurately due to their complex shapes and low-contrast boundaries with surrounding tissue. Automated segmentation must account for these dynamic patterns to avoid under-segmentation (*i.e.* false negatives). For visual illustration, we give in Figure 2.3 a malignant tumor seen in a DCE-MRI image taken from BreastDM dataset [211] presenting a complex shape and low contrast boundaries.

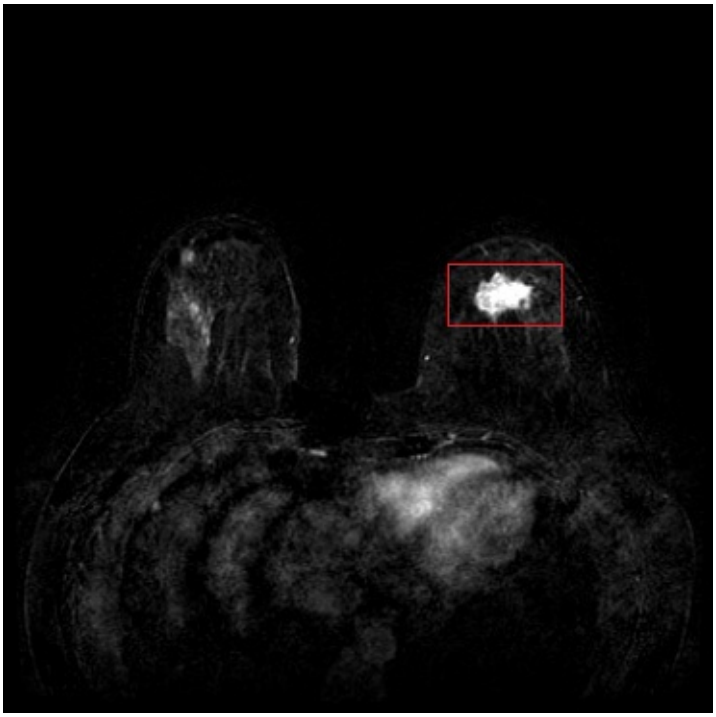

Figure 2.3: An illustration of a benign breast tumor highlighted with a bounded box in DCE-MRI taken from BreastDM dataset [211]

- **Normal (healthy) breast tissue:** It includes glandular and fatty components and usually shows minimal or no enhancement in DCE-MRI, with uniform, low-level signal intensity across phases [111]. However, physiological enhancement in healthy breasts, such as background parenchymal enhancement or veins enhancements, with their variability can imitate tumor signals. These latter require robust algorithms to avoid misclassifying healthy regions as tumors. Figure 2.4 shows enhanced heathy regions of the breast in a DCE-MRI image.

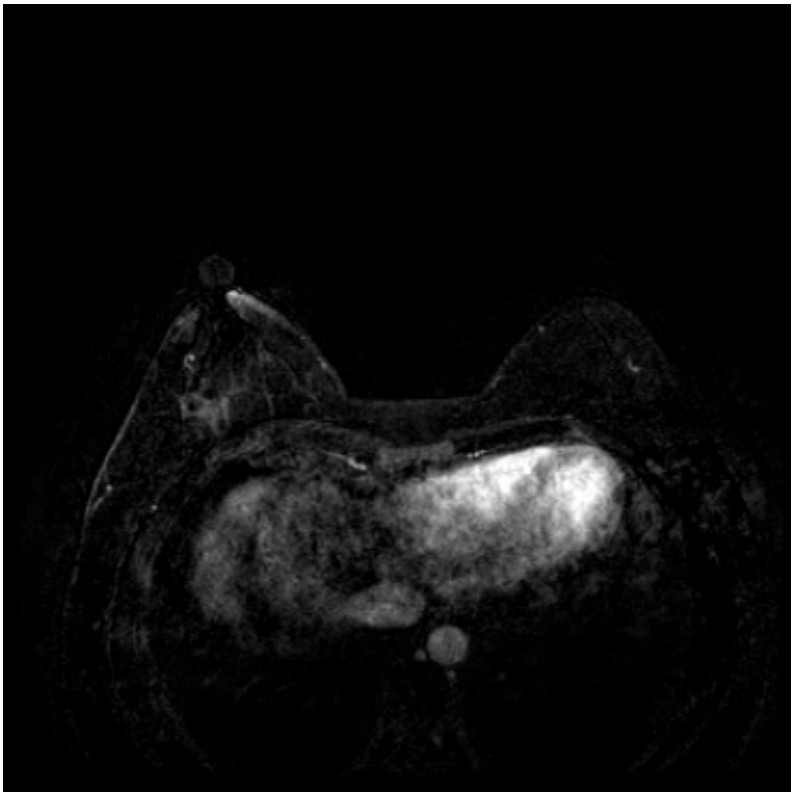

Figure 2.4: Illustration of a DCE-MRI image presenting different enhanced normal regions [211]

These latter exhibited cases present distinct challenges for segmentation methods. Especially for deep learning based models, such as U-Net, Vision Transformers, and TransUNet, that have to be trained on well-confined datasets to master as many cases as possible. Furthermore, other challenges like overlapping enhancement, motion artifacts, variations in acquisition process, and the size of tumors can badly affect segmentation methods [113].

2.3 Description of related work

Since the apparition of the first deep learning based models, several research projects have reported their successful application to medical image analysis. Furthermore, their exceptional performance compared to traditional handcrafting methods has driven computer vision scientists to continuously improve them in order to accomplish challenging tasks like tumor extraction from medical images. Hence, giving rise to a very specific fine-tuned deep learning-based approaches adapted to a particular task, dataset, or domain. DCE-MRI breast tumor segmentation is no exception to the fact, given the number of deep learning-based approaches proposed in its context over the past decade. As a main goal to stand out and discuss the essential progress made recently by researchers in the framework, we aim to review and succinctly discuss, in this section, some of the latest and relevant state-of-the-art methods presented in the setting. The analyzed approaches are

classified herein into *multiparametric* and *uniparametric data-driven methods*, based on the type of DCE-MRI data used for tumor feature extraction.

2.3.1 Multiparametric data-driven DCE-MRI breast tumor segmentation methods

These methods integrate multiple DCE-MRI sequences to extract comprehensive tumor features. Specifically, multiparametric data-driven approaches aim to enhance tumor representation by incorporating complementary information that may be missed when relying on a single sequence. However, these approaches require large datasets due to their dependence on multiple sequences. Building these datasets is challenging because of the complexity of the annotation and the limited availability of complete data in real-world clinical settings. Moreover, such methods present challenges when it comes to data heterogeneity and integration, as the sequences frequently differ in resolution, contrast, and noise characteristics. Furthermore, they are prone to overfitting and have a tendency to take too long to train, which can hinder their practical deployment. For instance, Peng et al. [153] have introduced an Inter-Modality Information Interaction Network (IMIIN), an advanced deep learning model designed to provide a 3D multi-modal breast tumor segmentation in MRI imaging. IMIIN employs a hierarchical two-stage approach that begins with a coarse tumor localization phase, followed by a fine-grained segmentation stage utilizing a 3D Tiny Object Segmentation Network (3D-TOSN). This hierarchical design effectively captures detailed boundary information, making it particularly adept at segmenting small or complex tumors that are often challenging for traditional models. A central innovation in IMIIN is its Bi-directional Request-Supply Information Interaction Module (BD-RSIIM), which facilitates a dynamic, mutual exchange of information between T1 and T1c MRI sequences. This bi-directional interaction allows each modality to selectively request and incorporate relevant information from the other, filtering out irrelevant data and thereby improving the accuracy and robustness of tumor boundary segmentation. While IMIIN demonstrates notable improvements over existing models, such as 2D-CMN and HDN, achieving superior sensitivity and Dice similarity scores, the model's dependence on well-aligned multi-modal data could limit its applicability in diverse clinical settings where imaging data may not be perfectly aligned. Additionally, the model's complex architecture and high computational requirements could pose challenges in resource-limited environments. To address these limitations, future enhancements could incorporate a pre-alignment module to handle unaligned data, broadening the model's utility across varied clinical workflows. Furthermore, exploring lightweight architectural designs could reduce computational demands, making IMIIN more feasible for deployment in settings with constrained resources. These refinements would enhance the model's versatility and make it more adaptable and practical for widespread clinical application in breast cancer diagnosis. Figure 2.5 presents the architecture of the IMIIN network as it is exhibited by Peng et al. [153].

Furthermore, Liu et al. [114] have presented a mask-guided 3D convolutional neural net-

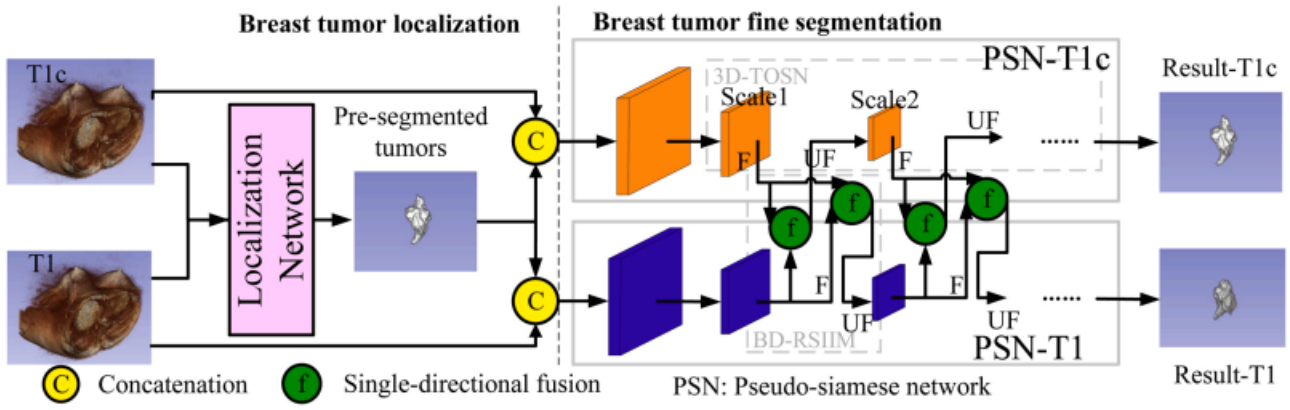

Figure 2.5: Illustration of the IMIIN network [153]

work (3D-CNN) that analysis both pre-contrast and post-contrast DCE-MRI sequences to predict breast cancer outcomes. Specifically focusing on five-year recurrence and HER2 status, by analyzing DCE-MRI images. The model introduces a mask-guided 3D-CNN architecture to direct the network's attention specifically to tumor regions, a method that enhances prediction accuracy over traditional models that either examine the entire image or focus solely on the tumor area. Techniques integral to this approach include dynamic contrast-enhanced MRI, which captures essential tumor characteristics, and class activation mapping (CAM), which visually highlights the regions that drive predictions, offering interpretability and precision. The model is built upon a modified 3D-VGGNet architecture optimized for volumetric data (see Figure 2.6 for illustration) and uses a loss function that combines cross-entropy with the Dice coefficient, promoting overlap between the model's attention and the tumor mask. To counter the small dataset (115 samples), 3D data augmentation techniques, such as rotations and translations, were employed, strengthening model robustness and reducing overfitting.

Despite promising results, there are limitations. Conservative segmentation might exclude relevant tumor areas, and the small dataset may limit generalizability. Future work could enhance this model by refining segmentation algorithms to include more of the tumor area, expanding the dataset, and comparing this mask-guided approach with traditional radiomics-based methods for further validation. This mask-guided 3D-CNN approach shows potential for improving the precision of prognostic predictions in clinical breast cancer applications, though further refinement and validation are essential.

Moreover, Zhang et al. [206] have presented a method for detecting and segmenting breast cancer in MRI using a Mask R-CNN model. Trained on non-fat-saturated images and tested on fat-saturated images. The model demonstrates impressive accuracy despite the differences between the training and testing data. It uses a data augmentation strategy that increases generalizability and robustness. The model achieves strong results in terms of Dice similarity coefficient and sensitivity when compared to existing methods. The approach of Zhang et al. could be helpful in clinical

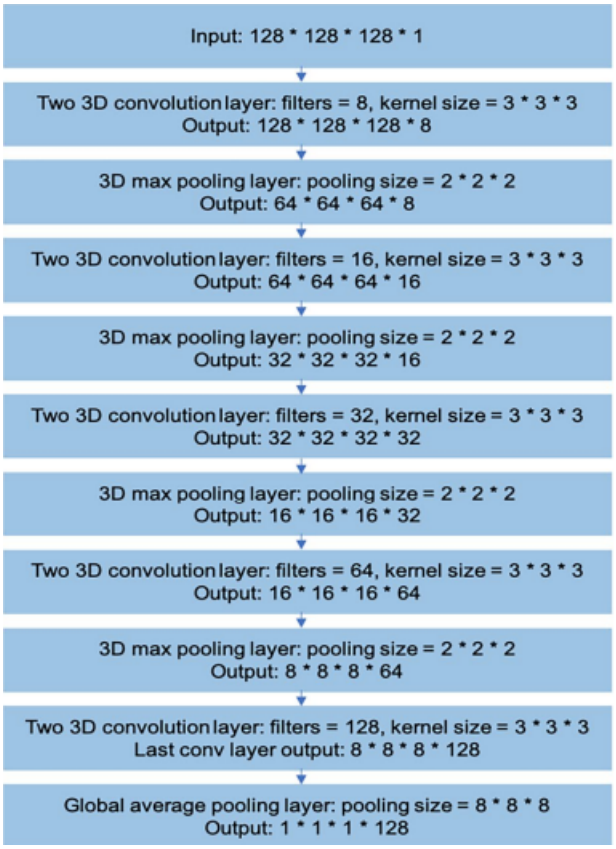

Figure 2.6: Modifed 3D VGGNet backbone architecture [114]

practice, offering a tool for automated breast cancer detection that could enhance radiologists' workflow and diagnosis accuracy. The detailed architecture of the Mask R-CNN model used by Zhang et al. [206] is illustrated in Figure 2.7.

2.3.2 Uniparametric data-driven DCE-MRI breast tumor segmentation methods

Within this category of methods, breast tumor segmentation is typically conducted using a single DCE-MRI sequence, most commonly the T1c sequence. These approaches have demonstrated computational simplicity when compared to multiparametric data-driven methods. Nevertheless, a single MRI sequence is often inadequate for capturing the complete morphological characteristics of the tumor under investigation. Furthermore, uniparametric data-driven methods tend to exhibit elevated false positive and false negative rates, primarily due to the resemblance between the enhancement patterns of tumorous and healthy tissues. For example, Hirsch et al. [77] have introduced a deep learning model based on a 3D U-Net convolutional neural network (CNN) to achieve fully automated breast tumor segmentation in MRI scans (see Figure 2.8 for illustration), reaching radiologist-level accuracy. The model was trained on an extensive dataset of over 60,000

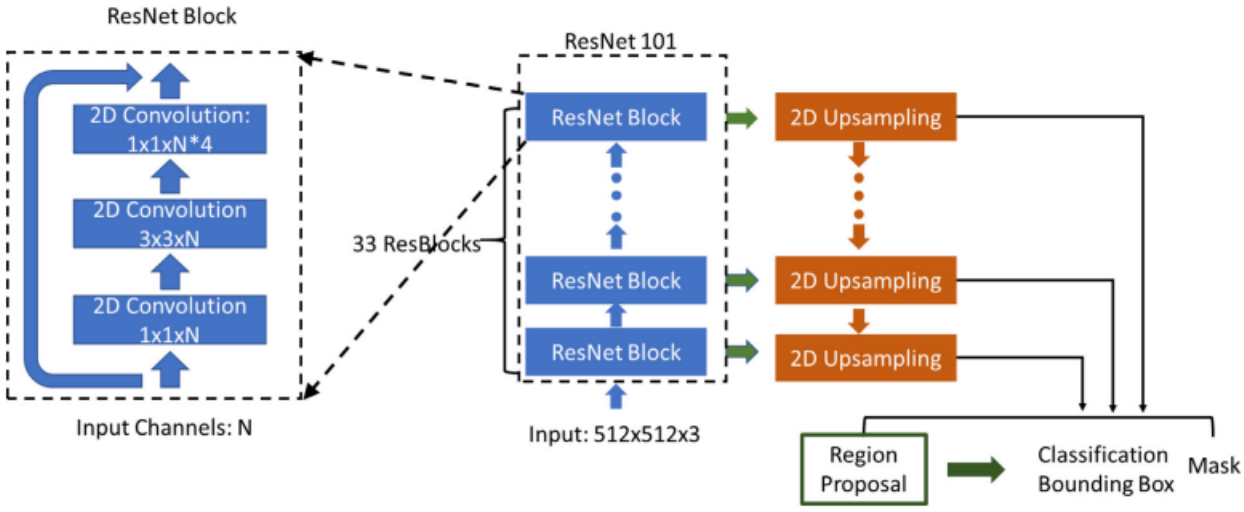

Figure 2.7: Mask R-CNN architecture used in the work of Zhang et al. [206]

clinical MRI scans, which included 2,555 segmented malignant cases and 60,108 benign cases, enabling robust feature extraction and superior generalization. By leveraging volumetric 3D imaging data rather than the conventional 2D slice-by-slice approach, the model captures complex spatial characteristics of tumors. Additionally, dynamic contrast enhancement (DCE) data and intensity normalization techniques were employed to harmonize imaging features across scans, enhancing segmentation precision. The model achieved a Dice score of 0.77, closely matching radiologists' performance (Dice scores of 0.69 to 0.84). Thus highlighting its potential as an efficient and accurate tool for breast MRI analysis.

Despite these advancements, the model has limitations in handling cases with specific tumor morphologies and dense breast tissue, particularly in patients with high background parenchymal enhancement (BPE), where segmentation accuracy decreases. The MRI protocol used (sagittal plane imaging) also limits direct application in settings favoring axial planes, which are more common in clinical practice. To address these issues, future work could involve training the model on larger, more diverse datasets that include high-resolution, multi-planar images, especially axial plane MRIs, and integrating further optimization techniques to improve accuracy in dense tissue cases. Expanding these capabilities could greatly enhance the model's robustness, clinical applicability, and diagnostic impact in breast imaging.

In addition, Rehman et al. [161] have proposed EEU-Net, an improved deep learning model for breast tumor segmentation in MRI scans, addressing the challenges of manual segmentation, tumor variability, and computational complexity. EEU-Net enhances the U-Net architecture by integrating Edge Encoder Blocks (EEB) and Edge Guidance Blocks (EGB) to focus on tumor boundaries, improving segmentation accuracy. The model was tested on the RIDER breast MRI dataset using five-fold cross-validation, achieving superior performance compared to state-of-the-art models, with a Dice coefficient of 0.884 and a Jaccard index of 0.804. Its custom loss function combines

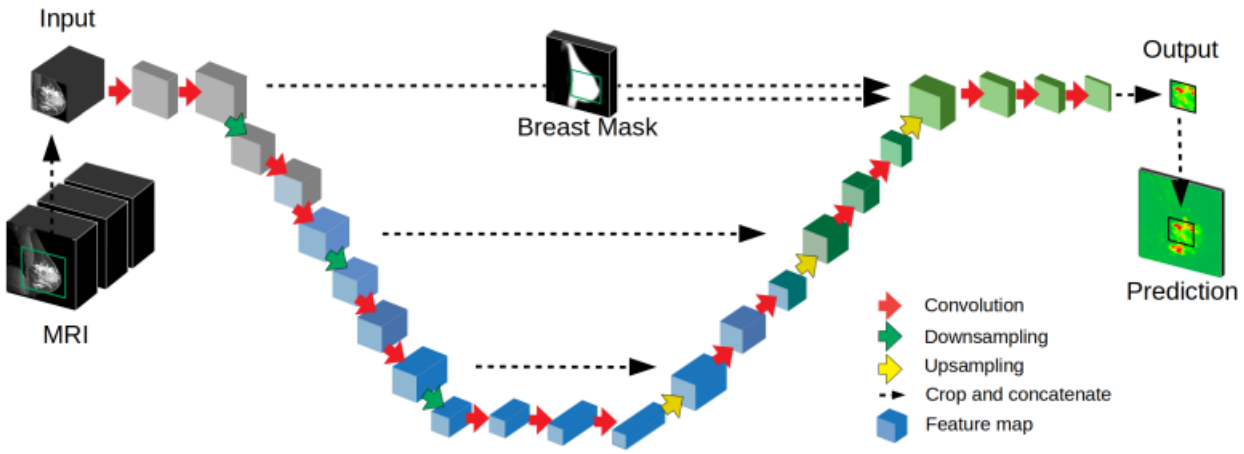

Figure 2.8: Illustration of the 3D U-Net convolutional neural network used for segmentation in the work of Hirsch et al. [77]

pixel loss and boundary loss, ensuring precise tumor region identification. Results demonstrate that EEU-Net effectively captures tumor structures and edges, outperforming traditional methods. Future work aims to apply the model to larger datasets, enhance computational efficiency, and incorporate multi-modal MRI inputs for further improvements in automated breast cancer diagnosis. The architecture of EEU-Net as it is proposed by Rehman et al. [161] is exhibited in Figure 2.9.

Furthermore, Wang et al. [187] have proposed a hybrid architecture combining 2D and 3D networks, using a technique called M2D3D-MC. The objective of this work is to combine the computational efficiency of 2D networks with the volumetric representation capabilities of 3D networks. The proposed method, M2D3D-MC, introduces three key contributions: (i) a hybrid 2D–3D convolutional module that addresses the loss of volumetric information inherent in 2D networks by integrating 3D spatial features extracted from adjacent slices; (ii) a multi-scale block designed to mitigate the impact of significant variations in lesion size on segmentation accuracy; and (iii) an image representation reconstruction module to enhance the overall segmentation performance. The M2D3D-MC model was evaluated on a dataset comprising 90 patients and 2,245 annotated axial slices. The method achieved promising results, with a Dice Similarity Coefficient (DSC) of 76.48, Sensitivity (SEN) of 75.93, and Positive Predictive Value (PPV) of 82.40. M2D3D-MC outperformed conventional segmentation models, including 2D U-Net, 3D U-Net, and other widely used approaches, demonstrating superior segmentation accuracy. Notably, the integration of a multi-scale block enabled more effective segmentation of lesions with varying sizes. Nevertheless, a key limitation of this study lies in the relatively small dataset, which may affect the model's generalizability. Expanding the dataset would likely enhance the robustness and overall performance of the proposed approach.

Moreover, Jiao et al. [68] have exhibited a 2D segmentation method based on U-Net++. The

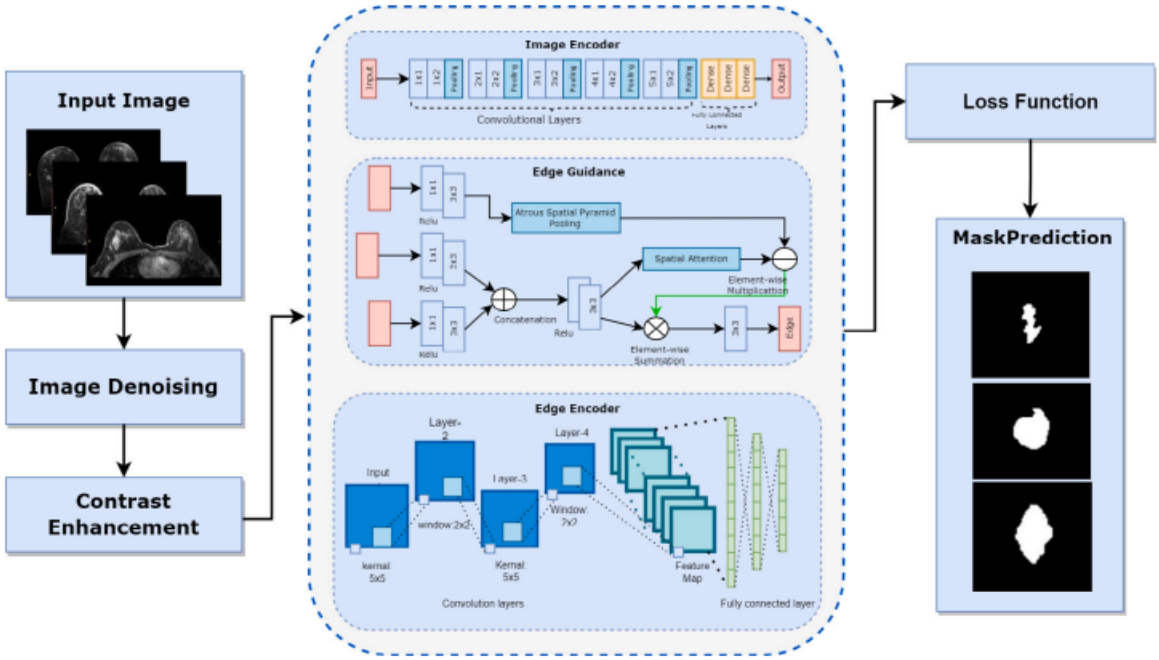

Figure 2.9: Architecture of EEU-Net proposed by Rehman et al. [161]

presented method is composed of two steps: in the first step, the whole breast region is segmented from the DEC-MRI image using the U-Net++. In the second step, Faster RCNN is employed to detect the mass like tumor from the produced segmented object of the breast region. The proposed segmentation technique was evaluated using a private dataset consisting of images from 75 patients and subsequently compared to U-Net. The results demonstrated a marginal improvement over U-Net, with performance metrics showing slight gains: Dice Coefficient of 0.951 (versus 0.941 for U-Net), Jaccard Index of 0.908 (versus 0.891 for U-Net), and Sensitivity of 0.948 (versus 0.941 for U-Net). May be this is due to the limited dataset size. Increasing the number of patients in the dataset could improve the generalizability of the method.

2.4 Comparison of the reviewed methods

Each of the reviewed deep learning-based methods exhibits distinct characteristics, each tailored to address specific aspects of tumor segmentation. In this section, we provide a comprehensive comparison of the analyzed studies based on four key criteria: the data dimensionality handled by each approach, the techniques employed, the dataset(s) used for training and evaluation, and the reported results which are presented as stated in the original published works. Specifically, the results are expressed in terms of three evaluation metrics, namely: *accuracy*, *Dice Similarity Coefficient (DSC)*, and *Jaccard index*. Mathematically, these evaluation metrics are formulated

using the parameters: *True Positives (TP)*, *True Negatives (TN)*, *False Positives (FP)*, and *False Negatives (FN)*, defined as follows:

- **TP**: refers to the number of correctly segmented contaminated pixels (voxels).
- **TN**: represents the number of correctly segmented uncontaminated pixels (voxels).
- **FP**: designates the number of uncontaminated pixels segmented as contaminated pixels (voxels).
- **FN**: is the number of contaminated pixels segmented as uncontaminated pixels (voxels).

The formulas for the evaluation metrics are given as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN},$$

it measures the overall correctness of the model by considering both correctly segmented contaminated and uncontaminated pixels (voxels).

$$\text{DSC} = \frac{2TP}{2TP + FP + FN},$$

it quantifies the overlap between predicted and ground truth segmentations, with values ranging from 0 (no overlap) to 1 (perfect overlap).

$$\text{Jaccard Index} = \frac{TP}{TP + FP + FN},$$

Also known as the Intersection over Union (IoU), it evaluates the similarity between predicted and actual segmentations.

Table 2.1 summarizes the key characteristics of these methods

2.5 Synthesis and insights

From Table 2.1, it is evident that most of the reviewed approaches employ full 3D deep learning models. Examples include Peng et al.'s IMIIN, which utilizes DenseVoxNet, as well as 3D-TOSN and BD-RSIIM. Additionally, Hirsch et al. propose a 3D U-Net–based method, while Zhang et al. implement Mask R-CNN, all of which leverage volumetric data to capture spatial context. In contrast, Liu et al. adopt a pseudo-3D approach, where a 2D CNN is enhanced with 3D contextual cues to predict prognostic outcomes. Meanwhile, Jiao et al. rely on a purely 2D architecture, specifically U-Net++, for segmentation and mass detection. This observation underscores the preference for 3D segmentation over 2D segmentation, as 3D representations provide a more comprehensive spatial perspective and enable the extraction of more discriminative features for tumor characterization.

Class	Work	Data dimension	Description	Technique(s)	Dataset(s)	Result(s)
Multiparametric data-driven methods	Peng et al. [153] (2022)	3D	Breast tumor segmentation in 3D multi-modal MRI using hierarchical networks and inter-modality information interaction	DenseVoxNet, 3D-TOSN, BD-RSIIM	Private dataset with 590 patients, 3D breast MRI scans (including T1c and T2c images)	DSC: 90.49% (T1c), 85.07% (T1)
	Liu et al. [114] (2021)	Pseudo-3D	Breast tumor prognostic outcome prediction on 3D DCE-MRI	Mask-Guided CNN (3D-VGGNet, CAM, mask-guided attention)	Private dataset with 115 patients (3D DCE-MRI scans)	Accuracy (training: 82.3%, validation: 79.6%)
	Zhang et al. [206] (2022)	3D	Automatic detection and segmentation of breast cancer in MRI	Mask R-CNN (Region-based CNN with Feature Pyramid Network, ResNet-101)	Private dataset with 241 patients (training) and 98 patients (testing). Non–fat-sat MRI for training, fat-sat MRI for testing	Per-slice detection accuracy of 86% (training) and 75% (testing). Per-lesion detection rate: 99.5% (training), 100% (testing). DSC : 82%) (training), 79%) (testing).
Uniparametric data-driven methods	Hirsch et al. [77] (2020)	3D	Fully automated breast tumor segmentation in MRI using deep learning	3D U-Net, DeepMedic, Conditional Random Field	Large private dataset: 2,555 malignant + 60,108 benign breasts for training and 250 test cases segmented by radiologists	DSC: 77%
	Rehman et al. [161] (2024)	2D	Boundary-driven deep learning for precise segmentation of breast tumors in DCE-MRI	EEU-Net	Public dataset: RIDER breast MRI dataset [**rider**]	Jaccard Index: 80% and DSC: 88%
	Wang et al. [187] (2021)	Mixed 2D and 3D	Lesion segmentation in breast DCE-MRI	M-2D3D	Private dataset with 90 patients, 2245 annotated axial slices	DSC: 76.48%, better segmentation of varying lesion sizes.
	Jiao et al. [68] (2020)	2D	Deep CNN-based automatic breast segmentation and mass detection in DCE-MRI	U-Net++	Private dataset with 75 patients	DSC: 95%, Jaccard Index: 0.91%

Table 2.1: Comparison of the reviewed methods according to few criteria.

Moreover, the majority of the reviewed methods were trained and evaluated on private datasets, which vary significantly in size and composition. Hirsch et al. utilized a large dataset comprising over 60,000 scans, facilitating robust training, whereas smaller datasets, such as Jiao et al.'s (75 patients), limit the generalizability of the model. Peng et al.'s dataset, consisting of 590 patients, and Zhang et al.'s 241 training cases provide a moderate-scale dataset; however, they lack standardization. In contrast, only one study leveraged a public dataset, specifically the RIDER breast MRI dataset. Consequently, comparing the performance of the analysed methods remains challenging due to variations in evaluation data. Despite these differences, the majority of approaches employ convolutional autoencoders with diverse architectures, highlighting their relevance in this domain.

Overall, the reviewed methods demonstrate that full 3D models excel in precise tumor delineation due to their ability to leverage volumetric context. In contrast, pseudo-3D and 2D methods can also achieve high effectiveness, particularly when enhanced with attention mechanisms or

multi-input strategies, making them suitable for both segmentation and prognostic tasks. The choice of architecture reflects a balance between dataset size, computational complexity, and specific clinical objectives, with each approach exhibiting distinct advantages in terms of accuracy, Dice Similarity Coefficient, Jaccard Index, and overall performance.

As the field of DCE-MRI breast tumor segmentation continues to advance, several promising research directions emerge. The following outlines key perspectives and potential strategies to further enhance this domain. Future efforts should focus on optimizing convolutional autoencoders, exploring alternative deep learning architectures, integrating multiple approaches, and developing novel models to achieve more accurate and efficient segmentation. The following are our suggestions for future works:

- All the reviewed methods were trained on well-aligned, high-quality datasets. To make these approaches more adaptable to real-world clinical settings, future research should focus on developing robust pre-alignment modules and domain adaptation techniques that can handle diverse datasets with varying MRI protocols and modalities.
- Full 3D models deliver great accuracy, but they come with high computational costs. To make them more practical for real-world clinical use, future research could focus on developing lighter architectures or using efficient training techniques like model pruning and knowledge distillation.
- Hybrid and pseudo-3D approaches hold great potential, but further research is needed to better capture 3D spatial relationships while keeping computational demands manageable.
- Although overall accuracy is high, segmenting small or irregularly shaped tumors remains a challenge. Integrating advanced loss functions, such as boundary-aware or focal loss, along with attention mechanisms designed to capture fine details, could help improve performance.
- The widespread use of private datasets limits reproducibility and comparability across studies. Creating and leveraging large, publicly available datasets would enable cross-validation, standardized benchmarking, and broader clinical adoption.

2.6 Conclusion

In conclusion, this chapter has presented a comprehensive review of DCE-MRI breast tumor segmentation research, focusing on recent state-of-the-art methods aimed at enhancing segmentation accuracy. Through a comparative analysis based on key criteria, we identified the strengths and limitations of existing approaches. This evaluation also outlined promising directions for future research to further advance the field and improve clinical outcomes.

In the next chapter, we will provide the necessary details and descriptions of our main contribution made in the setting of this work, by which we aim to further contribute to the progress in DCE-MRI breast tumor segmentation and ultimately improve patient outcomes.

Chapter 3

Empirical study of TransUNet and conditional GAN for DCE-MRI breast tumor segmentation

Introduction

This chapter details the complete methodological and experimental framework designed to investigate the efficacy of adversarial learning for breast tumor segmentation in DCE-MRI. Our central hypothesis, investigated in this chapter, is that while modern architectures like TransUNet (introduced by Chen et al. [31]) provide a strong baseline, their performance can be further enhanced by integrating a conditional Generative Adversarial Network (cGAN) framework. The adversarial loss, in theory, acts as a structural regularizer, compelling the generator to produce masks that are not only pixel-wise accurate but also anatomically plausible.

To rigorously verify this hypothesis, we conduct in this chapter a structured comparative study centered around three distinct model configurations. Our investigation begins with a foundational baseline: a *Standalone TransUNet* model trained in a purely supervised manner with a direct segmentation loss. This configuration serves to establish the performance ceiling of the TransUNet architecture under conventional training. Building on this baseline, we explore the impact of adversarial learning through two distinct cGAN frameworks. The first adversarial configuration pairs the TransUNet generator with a standard *PatchGAN discriminator* (proposed by Isola et al. [84]), designed to test the hypothesis that a critic focused on local high-frequency details can improve boundary sharpness. The second and more advanced configuration introduces a custom *hybrid CNN-Transformer discriminator*, designed to investigate whether a critic with access to local and global contextual information provides a more effective and stable training signal.

This structured comparative study allows us to isolate and evaluate the influence of adversarial supervision and discriminator design. The remainder of this chapter is organized as follows. Section 3.1 describes the BreastDM dataset and the data preparation pipeline. Section 3.2 provides

a detailed architectural breakdown of the generator and both discriminator models. Section 3.3 explains the loss functions used for both supervised and adversarial training. Section 3.4 outlines the training protocols and hyperparameters for each configuration. Section 3.5 defines the metrics used for quantitative assessment. Finally, Section 3.6 presents the complete experimental findings, including a critical discussion of the results.

3.1 Dataset and data preparation

A robust and well-prepared dataset is fundamental to the success of any deep learning model. This section details the data set used for this study and the pre-processing steps applied to ensure optimal model performance.

3.1.1 Dataset description

The experiments in this work involve the use of the BreastDM [210] dataset, a publicly available DCE-MRI collection specifically designed for breast tumor segmentation and classification. The BreastDM dataset was introduced in *Computers in Biology and Medicine* journal as a multi-institutional effort, comprising dynamic contrast-enhanced MRI scans acquired between January 2018 and September 2021 at Taizhou Central Hospital (Taizhou University, Zhejiang, China). The dataset includes 232 patients in total (147 with malignant tumors and 85 with benign tumors), with each undergoing three DCE-MRI phases, namely: pre-contrast, post-contrast, and subtraction sequences. All lesions were histologically confirmed and delineated by board-certified radiologists [210].

For the purposes of this work, we utilize the data from all 232 patients, encompassing both benign and malignant tumor cases. This approach allows us to train and evaluate our models on a more comprehensive and clinically representative dataset. Unlike studies that may concentrate on a single imaging phase, our methodology integrates multiple post-contrast and subtraction sequences to improve tumor representation. Specifically, we include five image subseries where available for each patient: VIBRANT+C8, SUB5, SUB6, SUB7, and SUB8. This multi-phase strategy aims to capture the full spectrum of enhancement characteristics across different tumor types, thereby offering a richer and more varied input distribution for training our models.

All DCE-MRI slices and their corresponding binary tumor masks were sourced from the official GitHub repository [175], which faithfully mirrors the data published by Zhao et al. [210]. The dataset was used with the predefined splits provided by the authors, and the slice-level alignment between images and masks was strictly preserved.

The final dataset statistics used in this study are as follows:

- **Training set:** 4,375 image–mask pairs (after filtering unmatched samples)

- **Validation set:** 585 image–mask pairs
- **Testing set:** 2,085 image–mask pairs

All images are stored in JPEG format as single-channel grayscale slices. Tumor masks are provided as binary PNG images, where tumor pixels are labeled with 1 and background pixels with 0.

3.1.2 Data preparation procedure

A comprehensive data preparation treatment was implemented using TensorFlow to process the raw DCE-MRI data into a format suitable for all three segmentation configurations. While the core preprocessing steps are consistent, slight variations exist depending on the model architecture and training strategy.

1. **File matching:** The procedure recursively scans the dataset directories to locate image files (JPEG) and corresponding tumor masks (PNG). A unique identifier is constructed for each sample by combining the patient ID and slice name to ensure strict one-to-one correspondence. Only valid image–mask pairs are retained.
2. **Resizing:** All images and masks are resized to a unified resolution of $\mathbf{256} \times \mathbf{256}$ pixels. Images are resized using bicubic interpolation to preserve grayscale fidelity, while masks are resized using nearest-neighbor interpolation to retain sharp boundary edges and prevent label distortion.
3. **Normalization:** Input images are standardized to single-channel (grayscale) and normalized from the original [0, 255] range to a floating-point range of $\mathbf{[0, 1]}$ to ensure numerical stability during training.
4. **Mask binarization:** All segmentation masks are binarized such that pixels with values above 0.5 are set to 1 (tumor), and the rest to 0 (background). This creates a clean and reliable target format for the binary segmentation task.
5. **Shuffling and batching:** All sub-datasets (training, validation, and test) are shuffled using a buffer size appropriate to the dataset split and batched into groups of size **8**. TensorFlow's `AUTOTUNE` functionality is used to optimize pipeline performance via prefetching.

For the baseline TransUNet model, additional data augmentation techniques, including: random flips and rotations, are applied during training using TensorFlow-native geometric transformations. These augmentations are intentionally omitted in the adversarial configurations (PatchGAN and Hybrid Discriminator) to maintain training stability in the GAN framework.

3.2 Model architectures and comparative framework

To investigate how different architectural components influence breast tumor segmentation performance, we implement in this study a comparative framework based on three progressively complex model configurations. All three models share a common generator backbone, the *TransUNet architecture* [31], but differ in their training paradigm and discriminator design. This setup allows us to isolate the impact of adversarial supervision and global context modeling through an ablation-style study.

The evaluated models are:

1. **Standalone TransUNet**: A purely supervised segmentation model serving as the baseline.
2. **TransUNet + PatchGAN (cGAN)**: A conditional GAN setup with a PatchGAN discriminator that enforces local realism in segmentation outputs.
3. **TransUNet + Hybrid CNN–Transformer Discriminator (cGAN)**: A more advanced conditional GAN model designed to capture both local and global structural coherence using a hybrid discriminator.

3.2.1 Generator's architecture: TransUNet

The generator (G) is common to all configurations and is based on TransUNet, a hybrid model that integrates Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs). This architecture benefits from CNNs' ability to model local textures and ViTs' strength in capturing long-range dependencies which represent a crucial aspect for accurately segmenting irregular breast tumors in DCE-MRI slices [49]. The TransUNet generator follows a three-stage architecture, which was previously detailed and illustrated in Chapter 1 (see Figure 1.13). To summarize its key components as implemented in our work:

- **CNN encoder**: A U-Net-like contracting path consisting of three convolutional blocks, each followed by max-pooling. It produces feature maps with channel depths of 64, 128, and 256, respectively. These features are later used in skip connections.
- **Transformer bottleneck**: At the network's deepest point, the 32×32 feature map is divided into non-overlapping patches and processed by a stack of four Transformer blocks. Each block contains Multi-Head Self-Attention and MLP layers, enabling the network to learn global contextual dependencies.
- **CNN decoder with skip connections**: The decoder upsamples the features back to the input resolution. At each stage, it concatenates the upsampled features with the corresponding encoder outputs via skip connections. The final output is produced by a 1×1 convolution followed by a sigmoid activation, generating a probability map for the tumor segmentation.

3.2.2 Discriminator's architectures for adversarial configurations

For the two adversarial configurations, we employ conditional Generative Adversarial Networks (cGANs), where the discriminator is conditioned on both the input image and the corresponding mask. This setup allows the discriminator to evaluate not only the realism of the predicted mask but also its plausibility in the context of the given DCE-MRI image [133]. Unlike classical GANs, which generate outputs based solely on random noise, cGANs leverage this conditioning to align the generation process with the spatial structure of the input. This characteristic makes cGANs especially well-suited for medical image segmentation tasks, where the anatomical consistency between the image and the mask is crucial for both clinical interpretability and model performance.

3.2.2.1 PatchGAN discriminator

The first adversarial setup employs a standard *PatchGAN* discriminator that has been exhibited by Isola et al. [84]. The PatchGAN discriminator is a fully convolutional network that outputs a matrix of real/fake predictions for overlapping image patches rather than the entire image.

- **Input**: A concatenated two-channel tensor composed of a 256×256 DCE-MRI slice and its corresponding binary mask.
- **Architecture**: Four strided convolutional layers followed by a $30 \times 30 \times 1$ output map, each value representing the realness of a patch.
- **Advantage**: Focuses on local high-frequency details, promoting sharper tumor boundaries in the generator output.

3.2.2.2 Hybrid CNN–Transformer discriminator

The second adversarial configuration uses a *hybrid discriminator* that combines CNN and Transformer components, designed to overcome PatchGAN's limitation in modeling global anatomical coherence.

- **Input**: Like PatchGAN, it processes a two-channel (image + mask) tensor.
- **CNN front-end**: Four convolutional layers downsample the input and extract a dense feature map.
- **Transformer back-end**: The CNN features are flattened into a sequence and passed through two Transformer blocks to capture global dependencies.
- **Output**: A single scalar prediction from a dense layer, representing the overall realness of the image–mask pair.

- **Advantage**: By modeling both local and global structure, it provides a stronger and more holistic training signal.

3.2.3 Training strategy connection

To ensure stability in adversarial training, we employ several recommended practices:

- **Asymmetric update frequency**: The discriminator is updated three times per generator update, preventing the generator from overpowering the discriminator early in training [120].
- **Label smoothing**: Real sample labels are smoothed to 0.9 instead of 1.0, a regularization technique known to stabilize GAN training and prevent overfitting [166].
- **Segmentation supervision**: In the both described cGANs, the generator is not only guided by adversarial loss but also supervised by a segmentation loss combining Binary Cross-Entropy and Dice Loss.

The next section provides a detailed description of these loss functions and optimization strategies used in each configuration.

3.3 Loss functions

This section outlines the loss functions employed across the three proposed configurations. Each model is optimized using losses appropriate to its architecture: a pure segmentation loss for the Standalone TransUNet and composite adversarial objectives for the two cGAN-based models (TransUNet + PatchGAN and TransUNet + Hybrid CNN–Transformer Discriminator).

3.3.1 Loss for the Standalone TransUNet

The standalone TransUNet is trained in a supervised setting using a composite segmentation loss that combines Binary Cross-Entropy (BCE) and Dice Loss. This design balances pixel-wise accuracy with structural similarity, especially in the presence of class imbalance is mathemathically expressed by Equation (3.1).

$$L_{\text{seg}} = \alpha \cdot L_{\text{BCE}} + (1 - \alpha) \cdot L_{\text{Dice}}, \tag{3.1}$$

where α is a weighting factor set to 0.2 in our experiments, prioritizing Dice Loss.

Analogously, the Binary Cross-Entropy Loss (L_{BCE}) is defined by Equation (3.2).

$$L_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^{N} \left[g_i \log(p_i) + (1 - g_i) \log(1 - p_i) \right], \tag{3.2}$$

and the Dice Loss (L_{Dice}) by Equation (3.3).

$$L_{\text{Dice}} = 1 - \frac{2 \cdot \sum_i g_i p_i + \epsilon}{\sum_i g_i + \sum_i p_i + \epsilon}, \tag{3.3}$$

where g_i and p_i represent the ground truth and predicted values at pixel i, respectively, and ϵ is a small smoothing constant to prevent division by zero.

3.3.2 Loss for cGAN-based models

For the adversarial configurations, the generator is trained to minimize both segmentation and adversarial losses, while the discriminator aims to distinguish real from generated masks. The total generator loss is defined by Equation (3.4):

$$L_G = L_{\text{seg}} + \lambda_{\text{adv}} \cdot L_{\text{adv}}, \tag{3.4}$$

where λ_{adv} is set to 0.01 to provide a mild adversarial signal while prioritizing accurate segmentation.

3.3.2.1 Generator Loss

The segmentation loss L_{seg} is the same as defined for the standalone TransUNet, combining BCE and Dice components. The adversarial loss L_{adv} varies slightly based on the discriminator design.

The Adversarial Loss for PatchGAN is expressed by Equation (3.5).

$$L_{\text{adv}} = L_{\text{BCE}}(D(x, G(x)), 1). \tag{3.5}$$

Specifically, the generator attempts to fool the discriminator into classifying fake segmentations as real (target = 1). Whareas, the Adversarial Loss for Hybrid Discriminator is defined by Equation (3.6).

$$L_{\text{adv}} = L_{\text{BCE}}(D_{\text{global}}(x, G(x)), 1). \tag{3.6}$$

Equation (3.6) captures both local and global inconsistencies by involving transformer-based context modeling in the discriminator.

3.3.2.2 Discriminator Loss

The discriminator in both cGAN setups is optimized to distinguish real image-mask pairs from fake ones. Its loss is given by Equation (3.7):

$$L_D = \frac{1}{2}(L_{\text{real}} + L_{\text{fake}}). \tag{3.7}$$

Explicitely, the Real Pair Loss with Label Smoothing L_{real} is given by Equation (3.8).

$$L_{\text{real}} = L_{\text{BCE}}(D(x, y), 0.9) \tag{3.8}$$

Label smoothing (target = 0.9) is applied to regularize learning and prevent overfitting [166].

On the other hand, the Fake Pair Loss L_{fake} is expressed by Equation (3.9).

$$L_{\text{fake}} = L_{\text{BCE}}(D(x, G(x)), 0). \tag{3.9}$$

Explicitly, the discriminator learns to assign a score close to 0 for synthetic outputs from the generator.

This adversarial training framework ensures that the generator not only segments tumors accurately but also produces masks that are realistic and anatomically consistent under scrutiny from a learned critic. The discriminator is trained three times more frequently than the generator to stabilize adversarial learning, and all loss values are monitored throughout training for convergence and evaluation.

3.4 Training strategy

The training protocols were carefully adapted for each of the three experimental configurations (*i.e* Standalone TransUNet (supervised baseline), TransUNet with PatchGAN discriminator, and TransUNet with a hybrid CNN–Transformer discriminator). Each setting employs a specific optimization setup, loss function, and training loop suitable for its architecture, ensuring fair comparison and stable convergence.

3.4.1 Optimization and hyperparameters setting

For the Standalone TransUNet, the model was trained as a standard segmentation network using the AdamW optimizer with a learning rate of 10^{-4} and weight decay of 10^{-5}. The loss function combined binary cross-entropy and Dice loss in a weighted ratio of 20% and 80%, respectively.

In the conditional GAN configurations, the generator and discriminator were trained using separate optimizers. The generator employed the Adam optimizer with a learning rate of 2×10^{-4}, while the discriminator used a slightly lower learning rate of 10^{-4} to prevent it from dominating the training dynamics. Both optimizers adopted $\beta_1 = 0.5$ and $\beta_2 = 0.999$, as recommended in the original DCGAN formulation [157]. To further stabilize adversarial training, gradient clipping was applied using a `clipnorm` of 1.0. Table 3.1 summarizes the hyperparameters used across all experiments.

Table 3.1: Training hyperparameters across all model configurations.

Hyperparameter	**Standalone TransUNet**	**GAN configurations**
Optimizer	AdamW	Adam
Generator Learning Rate	1×10^{-4}	2×10^{-4}
Discriminator Learning Rate	N/A	1×10^{-4}
Adam β_1	0.9	0.5
Adam β_2	0.999	0.999
Gradient Clipping (clipnorm)	No	1.0
Epochs	60	100
Batch Size	4–8	8

3.4.2 Training procedure

For the Standalone TransUNet model, training was conducted over 60 epochs using augmented training data. Validation was performed on non-augmented samples. Early stopping and learning rate reduction were implemented via `ReduceLROnPlateau` and `EarlyStopping` callbacks based on the validation Dice score. The model's weights were checkpointed at the epoch achieving the highest validation performance.

The adversarial training for both GAN variants followed a two-phase alternating schedule. Each training iteration within an epoch proceeded as follows:

- **Discriminator update:** The discriminator was trained for *three consecutive steps* per batch using real image–mask pairs and generated (fake) outputs from the current generator. Losses for real and fake predictions were averaged over these steps.
- **Generator update:** The generator was then updated once per batch, using the frozen discriminator to compute the adversarial loss component. Its total loss consisted of a segmentation loss (a weighted combination of BCE and Dice loss) and an adversarial loss encouraging realistic mask generation.

This asymmetric update scheme ensures that the discriminator provides a reliable gradient signal without dominating the training dynamics. Label smoothing (target value of 0.9 for real samples) was applied to further improve discriminator stability.

Model checkpoints were saved at the end of each epoch. Additionally, qualitative results were visualized every five epochs by plotting predicted segmentation masks alongside ground truth annotations. For the final evaluation, Test Time Augmentation (TTA) was employed on the test set to enhance robustness.

3.4.3 Monitoring and evaluation

Throughout training, all loss components (including segmentation, adversarial, and discriminator-specific losses) were logged per epoch for post-training analysis. For the standalone model, Dice score, binary accuracy, and mean IoU were used as evaluation metrics. For the adversarial configurations, the same metrics were applied, along with visual inspection of generator outputs to assess realism and boundary quality.

Evaluation on the test set was conducted using TTA in the standalone and PatchGAN settings, aggregating predictions from multiple augmentations to generate more robust segmentation maps. All experiments were conducted using TensorFlow on Google Colab Pro with GPU acceleration.

3.5 Evaluation metrics

To assess segmentation performance, we employ two widely used overlap-based metrics in medical image analysis: the Dice Similarity Coefficient (DSC) and the Intersection over Union (IoU). These metrics are particularly appropriate for tumor segmentation tasks, where the foreground (i.e., tumor regions) often constitutes a small fraction of the image, making them robust to class imbalance.

Prior to evaluation, the generator's probabilistic output is thresholded at 0.5 to produce binary masks (P), enabling direct comparison with the binary ground truth masks (G).

3.5.1 Dice Similarity Coefficient (DSC)

The Dice Similarity Coefficient (DSC) quantifies the degree of overlap between the predicted and ground truth regions. For the mathematical formula of DSC, the reader may refer to Chapter 2 Section 2.4.

Moreover, a DSC value of 1 indicates perfect overlap, whereas 0 indicates no overlap. This metric emphasizes accurate prediction of positive class pixels and is especially useful in the presence of class imbalance [132].

3.5.2 Intersection over Union (IoU)

Also known as the Jaccard Index, IoU provides a stricter measure of overlap. Its mathematical definition is given in Chapter 2 Section 2.4.

Like DSC, an IoU of 1 indicates a perfect match between prediction and ground truth. However, IoU penalizes both false positives and false negatives more severely than DSC, making it a more conservative evaluation metric.

3.5.3 Evaluation protocol

The evaluation is conducted across the training, validation, and test sets. For the test set, a simple test-time augmentation (TTA) strategy is applied, consisting of horizontal flipping. The final prediction is computed by averaging the outputs from the original and flipped inputs, which helps improve robustness to spatial variations.

All metrics are calculated on a per-batch basis by flattening both the predicted and ground truth masks. The individual scores are then averaged over the entire dataset to produce the final evaluation results. A small smoothing constant $\epsilon = 10^{-7}$ is included where necessary to ensure numerical stability, particularly in the denominator of ratio-based metrics.

This standardized evaluation protocol ensures consistent and fair comparisons across all model configurations considered in this study.

3.6 Experimental results and comparative evaluation

This section presents a comprehensive evaluation of the three model configurations developed for breast tumor segmentation in DCE-MRI. Each model was rigorously assessed using the same training and validation splits from the BreastDM dataset to ensure a fair and consistent comparison. The evaluation includes both quantitative analysis (based on standard segmentation metrics such as Dice coefficient and Intersection over Union (IoU)) and qualitative visual comparisons of the predicted segmentation masks.

This comparative study aims to elucidate the contributions of adversarial learning and discriminator architecture to segmentation performance, particularly in capturing fine tumor boundaries and maintaining anatomical coherence.

3.6.1 Quantitative evaluation

To quantitatively evaluate our framework, we conducted a comprehensive performance analysis on the held-out test set of the BreastDM dataset using the DSC and IoU as our primary metrics.

Our analysis begins with an internal ablation study to systematically assess the impact of our different design choices. We compared the performance of three model configurations:

- A **Standalone TransUNet** trained with only direct segmentation loss.
- A conditional GAN using a **PatchGAN discriminator**.
- A conditional GAN using a **hybrid CNN-Transformer discriminator**.

The results of this ablation study are detailed in Table 3.2.

Table 3.2: Ablation study results comparing our three implemented model configurations on the BreastDM test set. The mean and standard deviation are reported.

Method	**Mean Dice (DSC) [%]**	**Mean IoU [%]**
Standalone TransUNet	**74 ± 33.16**	**67 ± 32.40**
TransUNet (G) + PatchGAN (D)	64 ± 33.79	55 ± 31.96
TransUNet (G) + Hybrid Discriminator (D)	61 ± 34.88	52 ± 32.29

The results yield a critical insight: the Standalone TransUNet achieved the best performance, significantly outperforming both adversarially trained models. This suggests that for this specific dataset and training configuration, the direct supervision signal from the BCE and Dice loss was a more effective and stable optimization objective than the one provided by the adversarial framework.

To provide a more comprehensive performance overview, we report the segmentation results of our best-performing model (Standalone TransUNet) separately on benign and malignant cases from the BreastDM dataset, as shown in Table 3.3. This breakdown allows insight into how the model generalizes across different tumor types. Interestingly, the results presented in the table indicate that the standalone TransUNet achieved better performance on malignant tumor cases than on benign ones. This finding is somewhat unexpected, given that malignant tumors are widely recognized in the literature as being more challenging to segment due to their irregular shapes, internal heterogeneity, and ambiguous boundaries that often overlap with healthy breast tissues, particularly the parenchyma.

Table 3.3: Performance breakdown of the best-performing model (Standalone TransUNet) on benign and malignant test cases from the BreastDM dataset.

Metric \ **Pathology**	**Benign**	**Malignant**
Dice (DSC) [%]	68	72
IoU [%]	60	64

Finally, to contextualize our work, we compare the performance of our three proposed configurations against several established segmentation models evaluated on the same BreastDM dataset. The benchmark results for external methods are adapted from the *BC-MRI-SEG* benchmark paper presented by Bilic and Chen [13]. It is important to emphasize that only the reported performances of 2D methods were considered for comparison, ensuring a fair and consistent evaluation, as our models are designed to operate on 2D axial slices of breast DCE-MRI volumes.

Table 3.4: Performance comparison of our three model configurations against established methods on the BreastDM test set. Benchmark results for external methods are adapted from the BC-MRI-SEG paper [13].

Method	**Mean Dice (DSC) [%]**	**Mean IoU [%]**
Other methods (from Bilic and Chen [13])		
U-Net 2D [164] (2015)	49	/
U-Net 2.1D [13] (2024)	60	/
Swin UNETR [71] (2022)	**75**	/
Med-SA [190] (2023)	73	/
Our implemented configurations		
Standalone TransUNet	74	67
TransUNet + PatchGAN	64	55
TransUNet + Hybrid Discriminator	61	52

As shown in Table 3.4, the Standalone TransUNet achieved a Dice score of 74%, rivaling advanced models like Swin UNETR (75%) and Med-SA (73%), while maintaining a simpler architecture. This underscores its effectiveness in capturing complex tumor features. Additionally, the cGAN-based variants, though slightly lower in performance, demonstrate promising potential for enhancing structural detail through adversarial learning.

3.6.2 Qualitative analysis

To provide deeper insight into the behaviors of the implemented models, this section presents a qualitative analysis of segmentation results on representative cases from the validation set. We examine the model's performance in three key areas: the fidelity of tumor contours, robustness to morphological variation, and the handling of internal tumor heterogeneity.

3.6.2.1 High-fidelity contour segmentation

A primary advantage of the adversarial training framework is its ability to enforce structural realism. All three configurations studied (Standalone TransUNet, TransUNet + PatchGAN, and TransUNet + Hybrid Discriminator) exhibit this capability to varying degrees. Figure 3.1 presents three representative segmentation examples. Each row shows the input image, the ground truth mask, and the predicted mask for a specific case. These examples illustrate how each model captures sharp and accurate tumor boundaries, closely adhering to the ground truth.

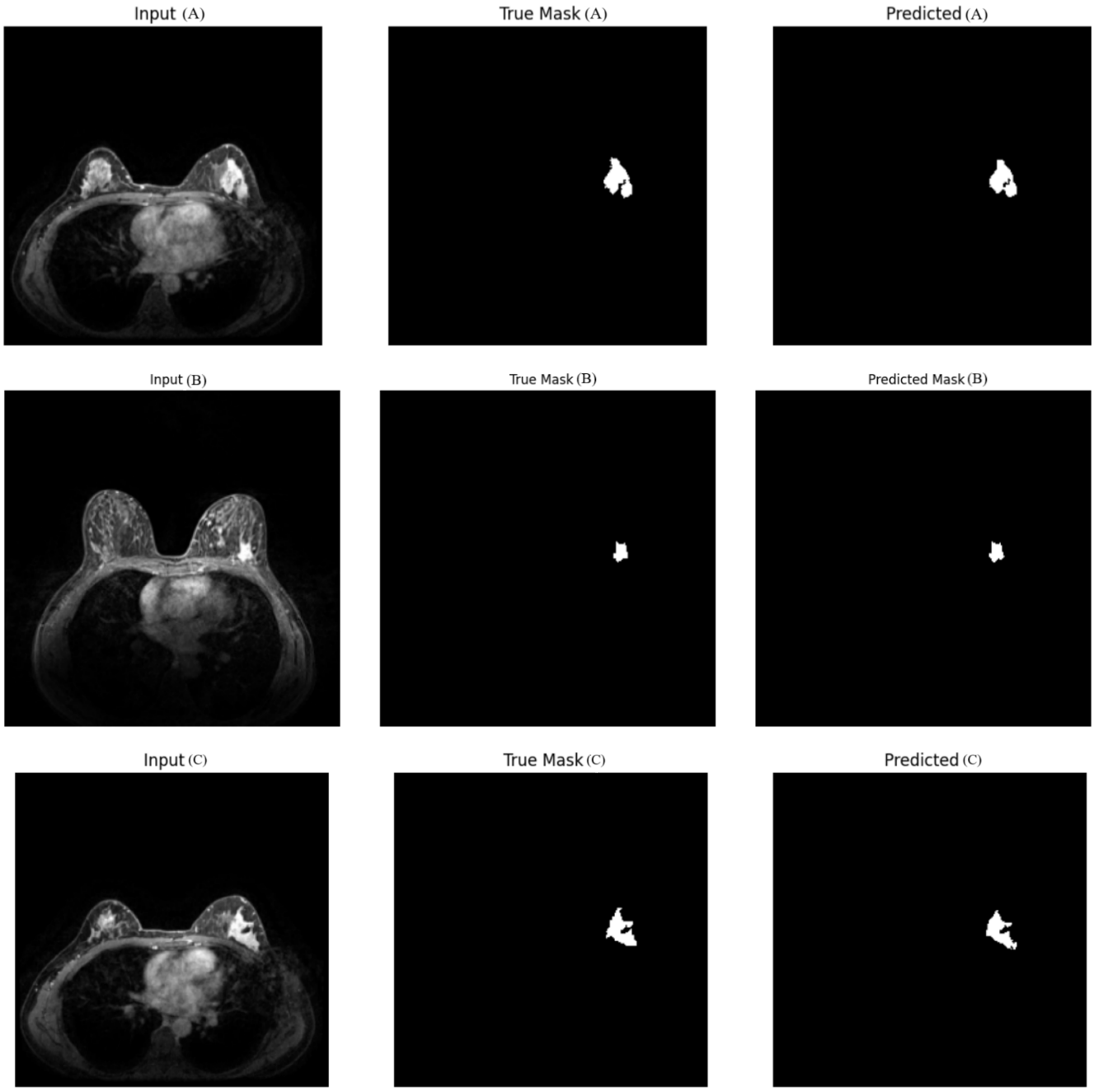

Figure 3.1: Qualitative results of high-fidelity contour segmentation. Each row shows input, ground truth, and predicted mask. **(A)** PatchGAN-based configuration; **(B)** Standalone TransUNet; **(C)** Hybrid CNN–Transformer discriminator.

3.6.2.2 Robustness to morphological variation

A critical measure of a clinical segmentation model's utility is its ability to perform reliably across the diverse range of tumor morphologies encountered in practice. The qualitative results presented in Figure 3.2 provide a necessary visual explanation for the quantitative findings reported in Table 3.2, showcasing the distinct failure modes and error patterns of each configuration.

A close inspection of the figure, particularly the regions highlighted by the *red boxes*, reveal examples of the variations between the predictions and the true masks. Interrestingly, the most significant observation is that the implemented models *exhibit a second, non-existent lesion in the ground truths*. This type of large-scale artifact suggests that either the models, while enforcing local realism, failed to maintain global anatomical consistency, leading to severe errors that drastically lower the quantitative scores. Or, the lesions do exist in the analyzed DCE-MRI images but erroneously characterized while establishing the ground truths. Therefore, a deep analyze of the yieled segmentations, involving experts in ragiology, is necessary to stand out the hallucinations from true tumoral cases.

Furthermore, it is important to analyze these predictions with an understanding of the inherent challenge in evaluating probabilistic models against binary ground truth, as suggested during our review. The models generate "soft" probabilistic maps, which may realistically capture the fuzzy, indistinct boundaries of tumors. The necessary conversion of these nuanced outputs into hard binary masks for metric calculation can *artificially penalize the models*, as they are measured against an idealized representation that may be an oversimplification of the underlying pathology. This mismatch between a realistic probabilistic output and a simplified binary target should be considered when interpreting the final scores; the true clinical utility of the predictions, despite the highlighted artifacts, may be higher than the quantitative metrics suggest.

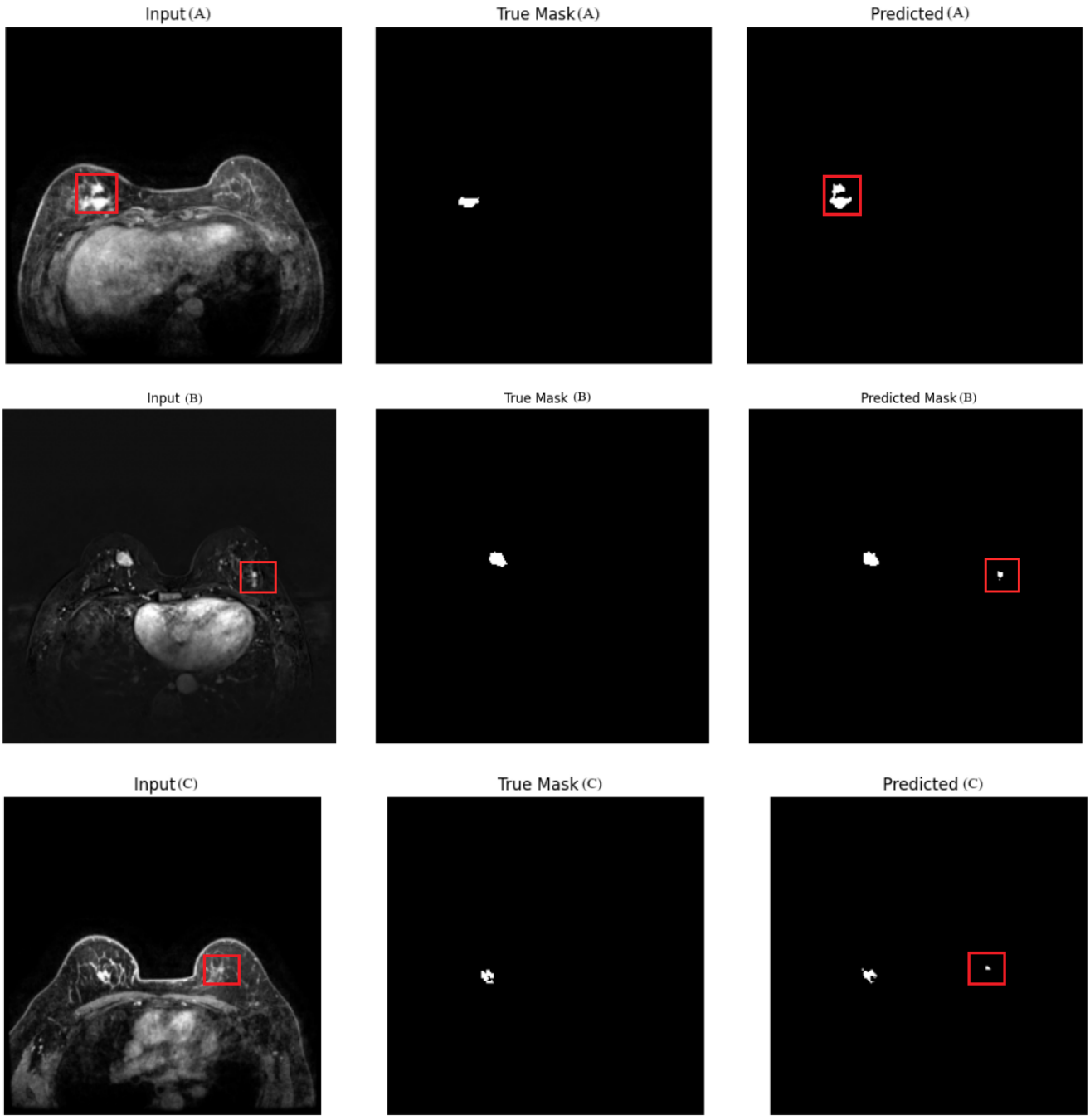

Figure 3.2: Qualitative results showing robustness to morphological variation. Each row displays the performance of a different model configuration on a distinct clinical case: **(A)** TransUNet + PatchGAN on a large, irregular tumor; **(B)** Standalone TransUNet on a compact lesion; **(C)** TransUNet + Hybrid Discriminator on another compact lesion.

3.6.2.3 Handling of internal heterogeneity

Malignant tumors often present with heterogeneous internal structures, such as necrotic cores, non-uniform enhancement patterns, or sub-regional texture variations. These internal differences can confuse segmentation models that rely heavily on intensity-based cues, leading to fragmented

or incomplete predictions. A robust segmentation model should learn the higher-level concept of a tumor and generalize across such internal noise. All three of our model configurations demonstrated strong capability in handling this heterogeneity as shwon in Figure 3.3 via the three representative examples. Despite internal visual variation, the predicted masks remain complete and contiguous, closely aligning with the ground truth.

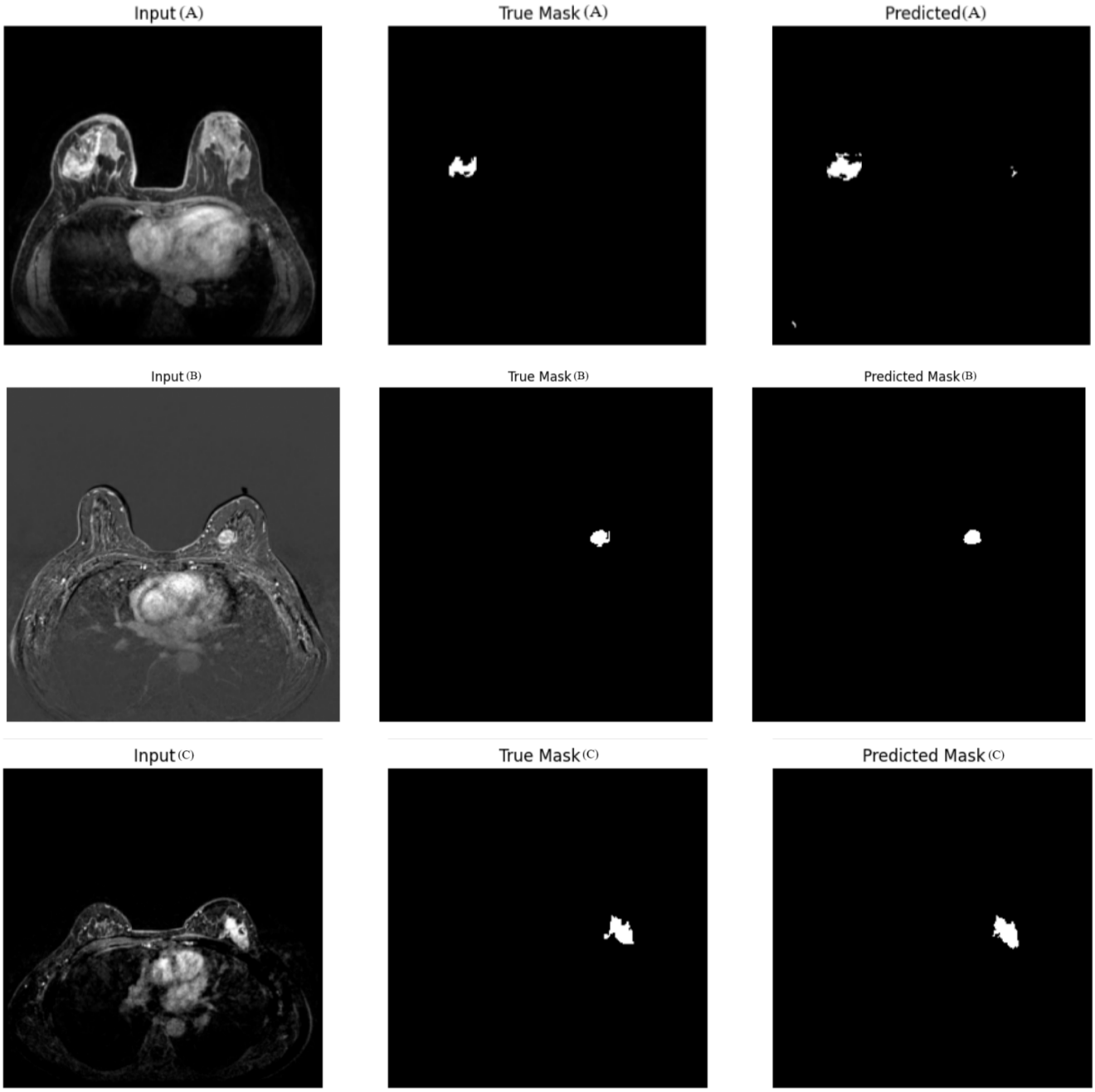

Figure 3.3: Qualitative results showing robustness to internal tumor heterogeneity. Each row displays the performance of a different model configuration: **(A)** TransUNet + PatchGAN on a lesion with necrotic regions; **(B)** Standalone TransUNet on a tumor with mixed intensities; **(C)** TransUNet + Hybrid Discriminator on a tumor with multi-textured internal variation.

3.6.3 Model convergence and generalization

To analyze the learning behavior and generalization capability of our models, we tracked their performance throughout the training process. We first examine the convergence of our best-performing model, the Standalone TransUNet, by plotting its segmentation loss and Dice score over 50 epochs (see Figure 3.4). The training loss shows a consistent and smooth decrease, while the training Dice score steadily increases, indicating that the model successfully learned the segmentation task on the training data. The validation curves, however, exhibit significant volatility. While the overall trend of the validation Dice score is positive, the high-frequency spikes suggest that the model's performance on unseen data is somewhat unstable, which aligns with the high standard deviation reported in our quantitative results (Table 3.2). Nonetheless, the lack of a persistent, widening gap between the training and validation curves indicates that the model is not severely overfitting.

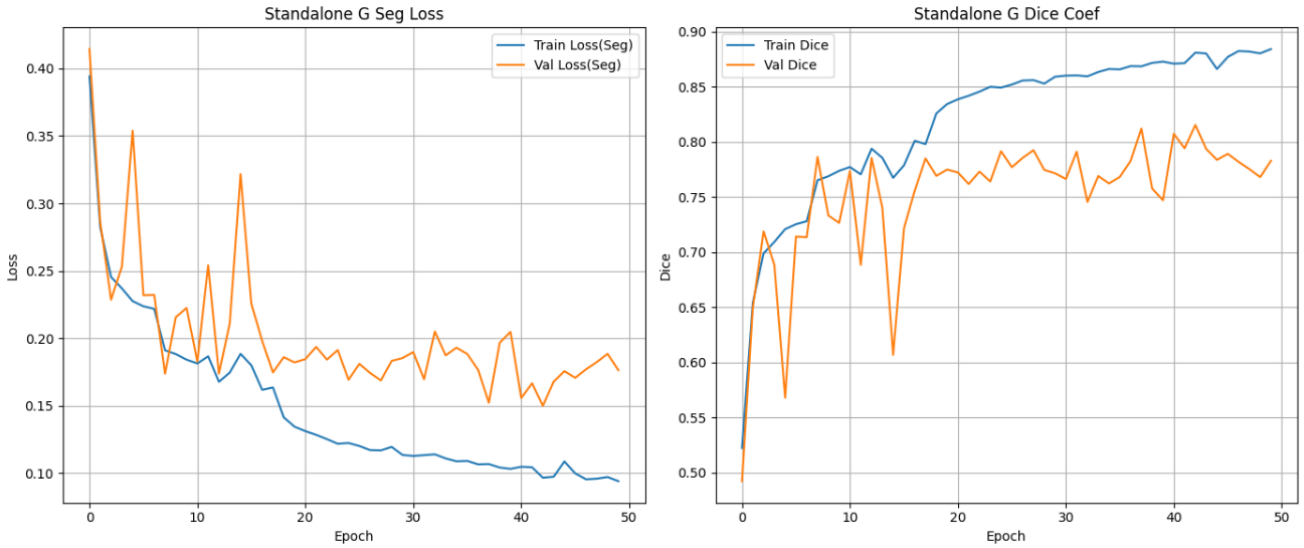

Figure 3.4: Training and validation curves for the Standalone TransUNet model. Left: Segmentation loss. Right: Dice Score. The smooth training curves and noisy validation curves highlight successful learning but with some instability on unseen data.

3.6.4 Adversarial training dynamics

To understand why the adversarially trained models underperformed, we now analyze their training dynamics. Figure 3.5 and Figure 3.6 show the loss curves for the TransUNet + PatchGAN and TransUNet + hybrid discriminator configurations, respectively.

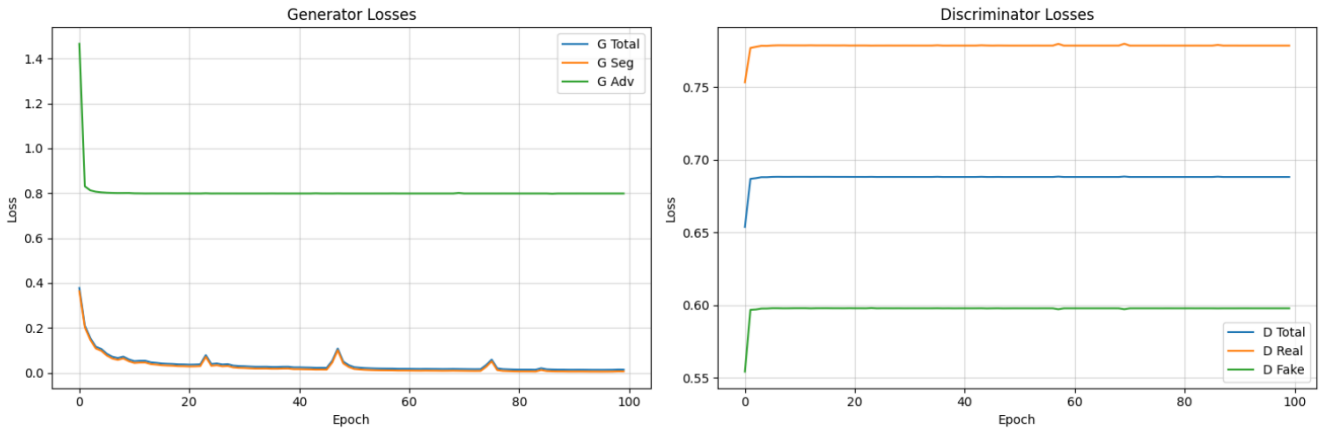

Figure 3.5: Loss curves for the TransUNet + PatchGAN model. Left: Generator losses. Right: Discriminator losses. The high, flat adversarial loss (G Adv) suggests the generator failed to fool the discriminator.

The TransUNet + PatchGAN framework (Figure 3.5) displays a deceptively stable training process. While the losses do not diverge, the dynamics reveal a dysfunctional equilibrium. The generator's segmentation loss ('G Seg') drops to a very low value, but its adversarial loss ('G Adv') remains high and flat. This indicates that the generator, while minimizing its pixel-wise error, completely failed to produce masks that could fool the discriminator. The discriminator's loss values are also non-ideal; they stabilize but at levels that suggest it easily distinguishes real from fake without providing a useful learning gradient back to the generator.

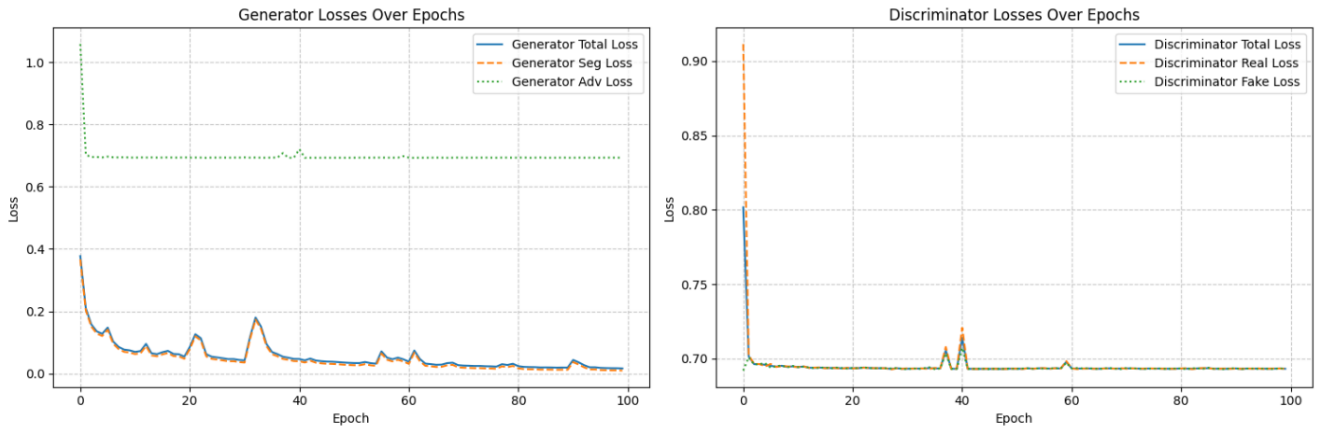

Figure 3.6: Loss curves for the TransUNet + hybrid discriminator model. The highly erratic and spiky nature of the discriminator loss (right) indicates severe training instability.

The TransUNet + hybrid Discriminator framework (Figure 3.6) exhibits even more problematic behavior. The discriminator loss plot is characterized by extreme instability, with large, erratic spikes throughout training. This suggests the hybrid discriminator struggled to find a stable decision boundary, providing noisy and inconsistent gradients. As with the PatchGAN model, the

generator's adversarial loss remains high, confirming that this configuration also failed to establish a productive adversarial game.

In summary, the loss dynamics for both GAN configurations confirm that the adversarial objective was counter-productive. The PatchGAN framework resulted in a "stuck" or unproductive equilibrium, while the hybrid discriminator framework led to severe training instability. This analysis provides a clear explanation for their poor quantitative performance compared to the stable and direct learning of the Standalone TransUNet.

3.7 Discussion

The comprehensive experimental results presented in this chapter lead to a clear, albeit counter-intuitive, primary conclusion: for the BreastDM dataset and our specific implementation, the addition of an adversarial training framework did not improve upon the performance of a well-tuned, standalone TransUNet model. This section interprets these findings, discusses their implications, and outlines the limitations of this study.

3.7.1 Interpretation of results

The quantitative analysis definitively established the Standalone TransUNet as the superior model, achieving a mean Dice score of 74%. This validates the inherent strength of the hybrid CNN-Transformer architecture, which effectively combines local feature extraction with global context modeling. The direct and unambiguous supervision signal from a combined BCE and Dice loss proved to be the most effective strategy for optimizing this model.

The most critical finding is the notable underperformance of both cGAN configurations. The analysis of their training dynamics (Subsection 3.6.4) provides a compelling explanation. The TransUNet + PatchGAN framework settled into a dysfunctional equilibrium where the generator, while minimizing its own segmentation loss, failed to produce masks that could meaningfully challenge the discriminator. The TransUNet + hybrid discriminator framework was even more problematic, suffering from severe gradient instability. In both cases, the adversarial objective acted as a **counter-productive regularizer**, forcing the generator to optimize for a conflicting goal of "realism" that ultimately detracted from the primary objective of precise anatomical alignment. While the qualitative outputs of the GANs may appear plausible at first glance, the lower Dice and IoU scores confirm a higher rate of subtle but significant errors in boundary placement.

3.7.2 Limitations of the current study

Despite the clear results, this work has several limitations that provide context for the findings and suggest directions for future work:

- **Adversarial framework sensitivity:** GAN training is notoriously sensitive to hyperparameter choices. It is possible that a more exhaustive search for the adversarial loss weight (λ_{adv}), different optimizer configurations, or more advanced stabilization techniques (e.g., spectral normalization, Wasserstein loss) could have yielded a more productive training equilibrium.

- **2D slice-level modeling:** The models operate on individual 2D slices, inherently ignoring the valuable 3D contextual information available across adjacent MRI slices. This limits the model's ability to ensure volumetric consistency.

- **Dataset characteristics:** While BreastDM is a valuable public resource, the size and variability of the dataset may still pose challenges for complex adversarial training, which often benefits from very large-scale data.

3.7.3 Directions for future work

Building on the findings and limitations of this study, several promising avenues for future research emerge:

- **3D architectures:** The most logical next step is to extend the framework to 3D by employing models like 3D U-Net or Swin UNETR to fully leverage volumetric data, which could significantly improve segmentation consistency and accuracy.

- **Advanced GAN techniques:** Instead of abandoning the adversarial approach, future work could investigate more robust GAN formulations (e.g., WGAN-GP) specifically designed to mitigate the training instabilities observed in this study.

- **Error analysis of the baseline:** A detailed analysis of the Standalone TransUNet's failure cases could reveal specific tumor morphologies or image characteristics where it struggles, guiding targeted architectural improvements or data augmentation strategies.

- **Multi-modal fusion:** Future models could incorporate other MRI sequences (e.g., T1-weighted, T2-weighted) or clinical metadata to provide a richer input representation and potentially improve performance.

3.8 Conclusion

This chapter presented a detailed empirical investigation into the use of a conditional Generative Adversarial Network (cGAN) framework for breast tumor segmentation. The central hypothesis was that an adversarial objective, implemented with either a PatchGAN or a hybrid CNN-Transformer discriminator, could regularize a powerful TransUNet generator to produce more accurate and realistic segmentation masks than a standalone, directly supervised model. To test this, we rigorously compared these three configurations on the public BreastDM dataset, using standard overlap-based metrics and qualitative analysis.

The principal finding of this chapter is that, contrary to our initial hypothesis, the addition of an adversarial training paradigm proved detrimental to segmentation performance. The Standalone TransUNet achieved the highest Dice and IoU scores, outperforming both cGAN variants by a significant margin. This quantitative underperformance was corroborated by the training dynamics, which revealed unstable loss patterns for the adversarial models, suggesting that the adversarial objective conflicted with, rather than complemented, the primary segmentation task. While qualitative analysis showed that all models could capture tumor contours with reasonable fidelity, the GAN variants did not produce visually superior masks that would justify their lower quantitative scores.

This outcome underscores a crucial observation in applied deep learning: advanced techniques like GANs are not universally beneficial and can introduce significant optimization challenges, especially in complex medical imaging tasks. The results suggest that the performance of these adversarial systems is highly sensitive to the balance of loss components, the stability of the discriminator, and potentially the size and variability of the training data. Therefore, the most valuable contribution of this chapter is the establishment of the standalone TransUNet as most effective and reliable baseline for this task, along with the clear empirical evidence of the pitfalls of this specific cGAN implementation. The insights gained here will directly inform the future directions of this research.

General conclusion

This manuscript has reported a comprehensive and rigorous investigation into breast tumor segmentation in dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI), emphasizing the integration and comparative evaluation of state-of-the-art deep learning architectures. Specifically, TransUNet and conditional Generative Adversarial Networks (cGAN). By leveraging the transformer-based self-attention capabilities of TransUNet alongside the adversarial learning strengths of cGANs, the study has demonstrated notable improvements in segmentation accuracy, boundary precision, and model robustness over conventional convolutional neural network approaches.

The empirical results obtained through extensive experimentation confirm that both TransUNet and cGAN frameworks, when applied individually or in combination, effectively capture complex tumor morphology and heterogeneous tissue characteristics inherent in breast MRI scans. TransUNet's ability to model long-range dependencies enhances global contextual understanding, while cGAN's adversarial training refines segmentation outputs by enforcing realistic tumor boundary delineation. This synergy contributes to more reliable and clinically relevant tumor segmentation, which is crucial for accurate diagnosis, treatment planning, and patient monitoring.

Despite these promising advances, the study also acknowledges persistent challenges, mainly intrinsic heterogeneity of breast tumors. These factors continue to limit the generalizability and scalability of automated segmentation models in real-world clinical applications. To address these issues, future research should explore hybrid architectures that further integrate transformer and adversarial learning paradigms, incorporate multi-modal imaging data to enrich feature representation, and develop advanced training strategies such as semi-supervised or self-supervised learning to mitigate annotation scarcity.

Overall, this thesis contributes significant insights and practical methodologies to the field of medical image analysis. It establishes a strong foundation for the development of more accurate, robust, and clinically applicable automated breast tumor segmentation systems, ultimately supporting improved patient outcomes and advancing the role of artificial intelligence in oncological imaging

Bibliography

[1] D. Abdelhafiz, C. Yang, R. Ammar, and S. Nabavi. "Deep Convolutional Neural Networks for Mammography: Advances, Challenges and Applications". In: *BMC Bioinformatics* 20 (2019), pp. 1–20.

[2] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. "SLIC Superpixels Compared to State-of-the-Art Superpixel Methods". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34 (2012), pp. 2274–2282.

[3] R. Adams and L. Bischof. "Seeded Region Growing". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16 (1994), pp. 641–647.

[4] S. C. Agner, J. Xu, and A. Madabhushi. "Spectral Embedding Based Active Contour (SEAC) for Lesion Segmentation on Breast Dynamic Contrast Enhanced Magnetic Resonance Imaging". In: *Medical Physics* 40 (2013), p. 032305.

[5] S. Amiri, M. M. Movahedi, K. Kazemi, and H. Parsaei. "An Automated MR Image Segmentation System Using Multi-Layer Perceptron Neural Network". In: *Journal of Biomedical Physics & Engineering* 3.4 (2013), pp. 115–122.

[6] S. S. Al-amri, N. V. Kalyankar, and S. D. Khamitkar. "Image Segmentation by Using Edge Detection". In: *International Journal on Computer Science and Engineering* 02 (2010), pp. 804–807.

[7] E.D.C. Anderson, B.B Muir, J.S. Walsh, and A.E. Kirkpatrick. "The efficacy of double reading mammograms in breast screening". In: *Clinical Radiology* 49 (1994), pp. 248–251.

[8] M. Arjovsky and L. Bottou. "Towards principled methods for training generative adversarial networks". In: *arXiv preprint arXiv:1701.04862* (2017).

[9] R. Azmi, R. Anbiaee, N. Norozi, L. Salehi, and A. Amirzadi. "A New Interactive Self-Training Approach to Segmentation Suspicious Lesions in Breast MRI". In: *Journal Medical Signals Sensors* 1 (2011), pp. 138–148.

[10] J. Besag. "On the Statistical Analysis of Dirty Pictures". In: *Journal of the Royal Statistical Society: Series B (Methodological)* 48 (1986), pp. 259–279.

[11] S. Beucher. "The Watershed Transformation Applied to Image Segmentation". In: *Scanning Microscopy* 28 (1992).

[12] R. Bilal, B. M. Khan, and R. Young. "Breast Cancer Detection and Diagnosis". In: *Emerging Developments and Practices in Oncology*. Ed. by I. El Naqa. 1st. Michigan: IGI Global, 2018, pp. 1–27.

[13] A. Bilic and C. Chen. "BC-MRI-SEG: A Breast Cancer MRI Tumor Segmentation Benchmark". In: *12th International Conference on Healthcare Informatics (ICHI)*. 2024, pp. 674–678.

[14] C. M. Bishop. "Training with Noise is Equivalent to Tikhonov Regularization". In: *Neural Computation* 7.1 (1995), pp. 108–116.

[15] F. Bloch and E. M. Purcell. "Nuclear Induction". In: *Physical Review* 70.7-8 (1946), pp. 460–474.

[16] F. Bouchebbah. "A New Levels Propagation Approach to Image Segmentation: Theory and Its Application to 2D/3D Breast MR Images". PhD thesis in Computer Science. University of Bejaia, July 2020.

[17] F. Bouchebbah and H. Slimani. "3D Automatic Levels Propagation Approach to Breast MRI Tumor Segmentation". In: *Expert Systems with Applications* 165 (2021), p. 113965. ISSN: 0957-4174.

[18] F. Bouchebbah and H. Slimani. "Levels Propagation Approach to Image Segmentation: Application to Breast MR Images". In: *Journal of Digital Imaging* 32 (2019), pp. 433–449.

[19] M. Boudiab and H. Bensassi. "Approche à Base d'Alliances dans les Graphes pour la Réduction de la Saturation et de la Congestion dans les Réseaux LTE". MA thesis. Département d'Informatique, Université de Béjaia, Juillet 2019.

[20] F. Bozkurt, C. Köse, and A. Sari. "An Inverse Approach for Automatic Segmentation of Carotid and Vertebral Arteries in CTA". In: *Expert Systems with Applications* 93 (2018), pp. 358–375.

[21] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal. "Global Cancer Statistics 2022: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries". In: *CA: A Cancer Journal for Clinicians* 72.3 (2022), pp. 209–249.

[22] R. C. Brigham, R. D. Dutton, T. W. Haynes, and S. T. Hedetniemi. "Powerful Alliances in Graphs". In: *Discrete Mathematics* 309 (2009), pp. 2140–2147.

[23] L. A. Brinton, M. M. Gaudet, and G. L. Gierach. *Cancer Epidemiology and Prevention*. 4th. New York: Oxford University Press, 2018, pp. 861–888.

[24] K. Broelemann, X. Jiang, and A. Schwering. "Automatic Understanding of Sketch Maps Using Context-Aware Classification". In: *Expert Systems with Applications* 45 (2016), pp. 195–207.

[25] K. Broelemann, X. Jiang, and A. Schwering. "Automatic understanding of sketch maps using context-aware classification". In: *Expert Systems with Applications* 45 (2016), pp. 195–207.

[26] K. S. Camilius and V. K. Govidan. "A Review on Graph Based Segmentation". In: *International Journal of Image* 5 (2012), pp. 1–13.

[27] J. Canny. "A Computational Approach to Edge Detection". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8 (1986), pp. 679–698.

[28] J. S. Cardoso and L. Corte-Real. "Toward a Generic Evaluation of Image Segmentation". In: *IEEE Transactions on Image Processing* 14 (2005), pp. 1772–1782.

[29] A. Chakravarty and J. Sivaswamy. "RACE-Net: A Recurrent Neural Network for Biomedical Image Segmentation". In: *Journal of Biomedical and Health Informatics* 23.3 (2018), pp. 1151–1162.

[30] J. Chen, Y. Lu, Q. Yu, et al. *TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation*. arXiv preprint arXiv:2102.04306. 2021.

[31] Jieneng Chen, Yongyi Lu, Qihang Yu, Xiangde Luo, Ehsan Adeli, Yan Wang, Le Lu, Alan L Yuille, and Yuyin Zhou. "TransUNet: Transformers make strong encoders for medical image segmentation". In: *arXiv preprint arXiv:2102.04306* (2021).

[32] Jieneng Chen, Jieru Mei, Xianhang Li, Yongyi Lu, Qihang Yu, Qingyue Wei, Xiangde Luo, Yutong Xie, Ehsan Adeli, Yan Wang, et al. "TransUNet: Rethinking the U-Net Architecture Design for Medical Image Segmentation Through the Lens of Transformers". In: *Medical Image Analysis* 97 (2024), p. 103280.

[33] P. C. Chen and T. Pavlidis. "Image Segmentation as an an Estimation Problem". In: *Computer Graphics and Image Processing* 12 (1980), pp. 153–172.

[34] W. Chen, M. L. Giger, and U. Bick. "A Fuzzy C-Means (FCM)-Based Approach for Computerized Segmentation of Breast Lesions in Dynamic Contrast-Enhanced MR Images". In: *Academic Radiology* 13 (2006), pp. 63–72.

[35] Y. Chen, S. Ma, X. Chen, and P. Ghamisi. "Hyperspectral Data Clustering Based on Density Analysis Ensemble". In: *Remote Sensing Letters* 8 (2017), pp. 194–203.

[36] S. R. Cherry and M. Dahlbom. *PET: Physics, Instrumentation, and Scanners*. 1st. New York: Springer, 2006.

[37] K. Chockley and E. Emanuel. "The end of radiology ? three threats to the future practice of radiology". In: *Journal of the American College of Radiology* 13 (2016), pp. 1415–1420.

[38] D. Cokuslu, K. Erciyes, and O. Dagdeviren. "A Dominating Set Based Clustering Algorithm for Mobile Ad Hoc Networks". In: *International Conference on Computational Science Berlin, Heidelberg*. Springer, 2006, pp. 571–578.

[39] *More Edge Detection (Web Resource)*. https://www.owlnet.rice.edu/~elec539/Projects97/morphjrks/moredge.html. (Accessed: 24/06/2025).

[40] T. Cover and P. Hart. "Nearest Neighbor Pattern Classification". In: *IEEE Transactions on Information Theory* 13 (1967), pp. 21–27.

[41] Y. Cui, Y. Tan, B. Zhao, L. Liberman, R. Parbhu, J. Kaplan, M. Theodoulou, C. Hudis, and L. H. Schwartz. "Malignant Lesion Segmentation in Contrast-Enhanced Breast MR Images Based on the Marker-Controlled Watershed". In: *Medical Physics* 36 (2009), pp. 4359–4369.

[42] E. Dam, M. Loog, and M. Letteboer. "Integrating Automatic and Interactive Brain Tumor Segmentation". In: *17th International Conference on Pattern Recognition*. IEEE, 2004, pp. 790–793.

[43] *Damadian's Invention (Image)*. https://commons.wikimedia.org/wiki/File:Damadian_invention.jpg. (Accessed : 24/06/2025).

[44] R. Damadian. "Tumor Detection by Nuclear Magnetic Resonance". In: *Science* 171 (1971), pp. 1151–1153.

[45] R. Damadian. "Tumor Detection by Nuclear Magnetic Resonance". In: *Science* 171.3976 (1971), pp. 1151–1153.

[46] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. "A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-II". In: *IEEE Transactions on Evolutionary Computation* 6 (2002), pp. 182–197.

[47] J. L. Deneubour, S. Aron, S. Goss, and J. M. Pasteels. "The Self-Organizing Exploratory Pattern of the Argentine Ant". In: *Journal of Insect Behavior* 3 (1990), pp. 159–168.

[48] A. Dosovitskiy, L. Beyer, A. Kolesnikov, et al. *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. arXiv preprint arXiv:2010.11929. 2020.

[49] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. "An image is worth 16x16 words: Transformers for image recognition at scale". In: *arXiv preprint arXiv:2010.11929* (2020).

[50] D. Dreizin, Y. Zhou, Y. Zhang, N. Tirada, and A. L. Yuille. "Performance of a Deep Learning Algorithm for Automated Segmentation and Quantification of Traumatic Pelvic Hematomas on CT". In: *Journal of Digital Imaging* (2019), pp. 1–9.

[51] M. El Adoui, S. A. Mahmoudi, M. A. Larhmam, and M. Benjelloun. "MRI Breast Tumor Segmentation Using Different Encoder and Decoder CNN Architectures". In: *Computers* 8 (2018), pp. 52–63.

[52] Mohammed El Adoui, Stylianos Drisis, and Mohammed Benjelloun. "Multi-Input Deep Learning Architecture for Predicting Breast Tumor Response to Chemotherapy Using Quantitative MR Images". In: *International Journal of Computer Assisted Radiology and Surgery* 15 (2020), pp. 1491–1500.

[53] W. Eziddin. "Segmentation Itérative d'Images par Propagation de Connaissances dans le Domaine Possibiliste: Application à la Détection de Tumeurs en Imagerie Mammographique". PhD thesis. Université Européenne de Bretagne, Brest, France, June 2012.

[54] H. Fan, C. Zhang, K. Zhang, and R. Yao. "An Improved CV Model for Breast MR Images Segmentation". In: *International Conference on Bioinformatics and Biomedicine*. IEEE, 2018, pp. 2243–2248.

[55] A. Q. Al-Faris, U. K. Ngah, N. A. M. Isa, and I. L. Shuaib. "Computer-Aided Segmentation System for Breast MRI Tumour Using Modified Automatic Seeded Region Growing". In: *Journal of Digital Imaging* 27 (2014), pp. 133–144.

[56] O. Favaron, G. Fricke, W. Goddard, S. M. Hedetniemi, S. T. Hedetniemi, P. Kristiansen, R. C. Laskar, and R. D. Skaggs. "Offensive Alliances in Graphs". In: *Discussiones Mathematicae Graph Theory* 24.2 (2004), pp. 263–275.

[57] P. F. Felzenszwalb and D. P. Huttenlocher. "Efficient Graph-Based Image Segmentation". In: *International Journal of Computer Vision* 59 (2004).

[58] E. Fix and J. L. Hodges Jr. "Discriminatory Analysis-Nonparametric Discrimination: Consistency Properties". In: *Technical Report, Randolf Field Texas: U.S. Air Force School of Aviation Medicine* (1951).

[59] G. W. Flake, S. Lawrence, and C. L. Giles. "Efficient Identification of Web Communities". In: *Proceedings of the 6th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2000)*. 2000, pp. 150–160.

[60] G. H. Fricke, L. M. Lawson, T. W. Haynes, S. M. Hedetniemi, and S. T. Hedetniemi. "A Note on Defensive Alliances in Graphs". In: *Bulletin of the Institute of Combinatorics and Its Applications* 38 (2003), pp. 37–41.

[61] C. Gallego-Ortiz and A. L. Martel. "A Graph-Based Lesion Characterization and Deep Embedding Approach for Improved Computer-Aided Diagnosis of Nonmass Breast MRI Lesions". In: *Medical Image Analysis* 51 (2019), pp. 116–124.

[62] C. Gallego-Ortiz and A.L. Martel. "A graph-based lesion characterization and deep embedding approach for improved computer-aided diagnosis of nonmass breast MRI lesions". In: *Medical Image Analysis* 51 (2019), pp. 116–124.

[63] A. Garcia-Garcia, S. Orts-Escolano, S. Oprea, V. Villena-Martinez, P. Martinez-Gonzalez, and J. Garcia-Rodriguez. "A Survey on Deep Learning Techniques for Image and Video Semantic Segmentation". In: *Applied Soft Computing* 70 (2018), pp. 41–65.

[64] P. Ghamisi, M. S. Couceiro, F. M. L. Martins, and J. A. Benediktsson. "Multilevel Image Segmentation Based on Fractional-Order Darwinian Particle Swarm Optimization". In: *IEEE Transactions on Geoscience and Remote Sensing* 52 (2013), pp. 2382–2394.

[65] https://gco.iarc.fr. (Accessed : 24/06/2025).

[66] A. B. de Gonzalez and S. Darby. "Risk of Cancer from Diagnostic X-rays: Estimates for the UK and 14 Other Countries". In: *The Lancet* 363 (2004), pp. 345–351.

[67] J. Gordillo, E. Montseny, and S. Sobrevilla. "State of the Art Survey on MRI Brain Tumor Segmentation". In: *Magnetic Resonance Imaging* 31 (2013), pp. 1426–1438.

[68] and X. Jiang H. Jiao, Z. Pang, X. Lin, Y. Huang, and L. Li. "Deep Convolutional Neural Networks-Based Automatic Breast Segmentation and Mass Detection in DCE-MRI". In: *Computational and Mathematical Methods in Medicine* 2020.1 (2020), p. 2413706.

[69] M. Ham and A. Moon. "Inflammatory and Microenvironmental Factors Involved in Breast Cancer Progression". In: *Archives of Pharmacal Research* 36 (2013), pp. 1419–1431.

[70] A. E. Hassanien, H. M. Moftah, A. T. Azar, and M. Shoman. "MRI Breast Cancer Diagnosis Hybrid Approach Using Adaptive Ant-Based Segmentation and Multilayer Perceptron Neural Networks Classifier". In: *Applied Soft Computing* 14 (2014), pp. 62–71.

[71] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. R. Roth, and D. Xu. "Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images". In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2021*. Springer, 2022, pp. 272–284.

[72] T. W. Haynes, S. T. Hedetniemi, and M. A. Henning. "Global Defensive Alliances in Graphs". In: *Electronic Journal of Combinatorics* 10 (2003), R47.

[73] K. He, X. Zhang, S. Ren, and J. Sun. "Deep Residual Learning for Image Recognition". In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2016.

[74] T. He, J. Song, Y. Zhu, J. Wang, G. Huang, Y. Qiao, and J. Sun. "Parametric Noise Injection: Trainable Randomness to Improve Deep Neural Network Robustness Against Adversarial Attack". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2020, pp. 588–597.

[75] D. Hendrycks, M. Mazeika, S. Kadavath, and D. Song. "Many Faces of Robustness: A Critical Analysis of Out-of-Distribution Generalization". In: *International Conference on Machine Learning (ICML)*. 2020, pp. 2751–2769.

[76] T. Herman. *Fundamentals of Computerized Tomography: Image Reconstruction from Projections*. 2nd. London: Springer Science and Business Media, 2009.

[77] Lukas Hirsch, Yu Huang, Shaojun Luo, Carolina Rossi Saccarelli, Roberto Lo Gullo, Isaac Daimiel Naranjo, A. G. Bitencourt, Natsuko Onishi, Eun Sook Ko, Dortis Leithner, et al. "Deep Learning Achieves Radiologist-Level Performance of Tumor Segmentation in Breast MRI". In: *ArXiv:2009.09827 Phys. Stat.* (2020).

[78] M. Hmida, K. Hamrouni, B. Solaiman, and S. Boussetta. "Mammographic mass segmentation using fuzzy contours". In: *Computer Methods and Programs in Biomedicine* 164 (2018), pp. 131–142.

[79] S. L. Horowitz and T. Pavlidis. "Picture Segmentation by a Tree Traversal Algorithm". In: *Journal of the ACM* 23 (1976), pp. 368–388.

[80] B. Huang, M. Warner, and J. A. Gustafsson. "Estrogen Receptors in Breast Carcinogenesis and Endocrine Therapy". In: *Molecular and Cellular Endocrinology* (2015), pp. 240–244.

[81] B. Huang, M. Warner, and J. A. Gustafsson. "Estrogen Receptors in Breast Carcinogenesis and Endocrine Therapy". In: *Molecular and Cellular Endocrinology* 418 (2015), pp. 240–244.

[82] K. M. Iftekharuddin, J. Zheng, M. A. Islam, and R. J. Ogg. "Fractal-Based Brain Tumor Detection in Multimodal MRI". In: *Applied Mathematics and Computation* 207 (2009), pp. 23–41.

[83] *IRM Diffusion (Vidéo)*. https://www.cea.fr/multimedia/pages/videos/culture-scientifique/sante-sciences-du-vivant/video-irm-diffusion.aspx. (Accessed : 24/06/2025).

[84] P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros. "Image-to-image translation with conditional adversarial networks". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 1125–1134.

[85] A. Israel. "Breast Cancer: The Past, the Present, and a Look Towards the Future". PhD thesis in Biology. University of Yeshiva, New York, June 2019.

[86] R. Jain and R. S. Sharma. "Image Segmentation Through Fuzzy Clustering: A Survey". In: *Harmony Search and Nature Inspired Optimization Algorithm*. Ed. by N. Yadav, A. Yadav, J. C. Bansal, K. Deep, and J. H. Kim. 1st. Singapore: Springer, 2019, pp. 497–508.

[87] J. Jayender, S. Chikarmane, F. A. Jolesz, and E. Gombos. "Automatic Segmentation of Invasive Breast Carcinomas from Dynamic Contrast-Enhanced MRI Using Time Series Analysis". In: *Journal of Magnetic Resonance Imaging* 79.5 (2014), pp. 467–475.

[88] J. S. Jeyanathan, A. Shenbagavalli, B. Venkatraman, and M. Menaka. "Analysis of Breast Thermograms in Lateral Views Using Texture Features". In: *IEEE TENCON 2018-2018 IEEE Region 10 Conference*. 2018, pp. 2017–2022.

[89] F. Jie, Y. Shi, Y. Li, and Z. Liu. "Interactive Region-Based MRF Image Segmentation". In: *4th International Congress on Image and Signal Processing - IEEE*. 2011, pp. 1263–1267.

[90] C. T. Johnston, K. T. Gribbon, and D. G. Bailey. "Implementing Image Processing Algorithms on FPGAs". In: *11th Electronics New Zealand Conference*. 2004, pp. 118–123.

[91] *Introduction to Image Segmentation with K-Means Clustering*. https://www.kdnuggets.com/2019/08/introduction-image-segmentation-k-means-clustering.html. (Accessed : 24/06/2025).

[92] M. Kass, A. Witkin, and D. Terzopoulos. "Snakes: Active Contour Models". In: *International Journal of Computer Vision* 1 (1988), pp. 321–331.

[93] H. Khastavaneh and H. Haron. "A Conceptual Model for Segmentation of Multiple Scleroses Lesions in Magnetic Resonance Images Using Massive Training Artificial Neural Network". In: *5th International Conference on Intelligent Systems, Modelling and Simulation*. IEEE, 2014, pp. 273–278.

[94] D. Kim, M. Park, J. Lee, H. B. Kim, and H. Kim. "Structured Noise Injection for Deep Learning Regularization". In: *Proceedings of the 30th ACM International Conference on Information & Knowledge Management (CIKM)*. 2021, pp. 1645–1654.

[95] A. Krizhevsky, I. Sutskever, and G. E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems*. 2012, pp. 1097–1105.

[96] C. K. Kuhl, H. B. Bieling, J. Gieseke, B. P. Kreft, T. Sommer, G. Lutterbey, and H. H. Schild. "Healthy Premenopausal Breast Parenchyma in Dynamic Contrast-Enhanced MR Imaging of the Breast: Normal Contrast Medium Enhancement and Cyclical-Phase Dependency". In: *Radiology* 215.1 (2000), pp. 153–161.

[97] C. K. Kuhl, H. B. Bieling, J. Gieseke, B. P. Kreft, T. Sommer, G. Lutterbey, and H. H. Schild. "Healthy Premenopausal Breast Parenchyma in Dynamic Contrast-Enhanced MR Imaging of the Breast: Normal Contrast Medium Enhancement and Cyclical-Phase Dependency". In: *Radiology* 215.1 (2000), pp. 153–161.

[98] C. K. Kuhl, S. Schrading, C. C. Leutner, N. Morakkabati-Spitz, E. Wardelmann, R. Fimmers, W. Kuhn, and H. H. Schild. "Mammography, Breast Ultrasound, and Magnetic Resonance Imaging for Surveillance of Women at High Familial Risk for Breast Cancer". In: *Journal of Clinical Oncology* 23.33 (2005), pp. 8469–8476.

[99] K. C. Kuhl, S. Schrading, C. Leutner, N. Morakkabati-Spitz, E. Wardelmann, R. Fimmers, W. Kuhnr, and H. H. Schild. "Mammography, Breast Ultrasound, and Magnetic Resonance Imaging for Surveillance of Women at High Familial Risk for Breast Cancer". In: *Journal of Clinical Oncology* 23.33 (2005), pp. 8469–8476.

[100] P. C. Lauterbur. "Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance". In: *Nature* 242 (1973), pp. 190–191.

[101] P. C. Lauterbur. "Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance". In: *Nature* 242.5394 (1973), pp. 190–191.

[102] P. C. Lauterbur. "Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance". In: *Nature* 242.5394 (1973), pp. 190–191.

[103] Y. LeCun, Y. Bengio, and G. Hinton. "Deep Learning". In: *Nature* 521.7553 (2015), pp. 436–444.

[104] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. "Gradient-Based Learning Applied to Document Recognition". In: *Proceedings of the IEEE* 86.11 (1998), pp. 2278–2324.

[105] B. D. Lee, K. C. Liu, S. F. Hung, P. C. Lei, T. L. Wang, and Yang. "Breast Tumor Classification of Ultrasound Images Using Wavelet-Based Channel Energy and Images". In: *Journal of Selected Topics in Signal Processing* 3 (2009), pp. 81–93.

[106] S. Lee, Q. Huang, L. Jin, M. Lu, and T. Wang. "A Graph-Based Segmentation Method for Breast Tumors in Ultrasound Images". In: *4th International Conference on Bioinformatics and Biomedical Engineering*. IEEE, 2010, pp. 1–4.

[107] I. Leichter, Z. Gallimidi, A. Heyman-Reiss, N. Merlet, E. Ratner, M. Abramov, I. Stainvas, and R. Lederman. "Is CAD able to assist in the detection of subtle breast findings-lobular cancers, and t1a/t1b masses in dense breasts?" In: *International Workshop on Digital Mammography*. Springer. 2010, pp. 474–480.

[108] D. Levi, N. Ben-Zrihem, and E. Fetaya. "Structured Noise Injection for Regularization in Deep Neural Networks". In: *International Conference on Learning Representations (ICLR)*. 2021.

[109] C. Li, R. Huang, Z. Ding, J. Gatenby, D. N. Metaxas, and J. C. Gore. "A Level Set Method for Image Segmentation in the Presence of Intensity Inhomogeneities with Application to MRI". In: *IEEE Transactions on Image Processing* 20 (2011), pp. 2007–2016.

[110] G. Litjens, T. Kooi, B. E. Bejnordi, et al. "A Survey on Deep Learning in Medical Image Analysis". In: *Medical Image Analysis* 42 (2017), pp. 60–88.

[111] G. Litjens, T. Kooi, B. E. Bejnordi, et al. "A Survey on Deep Learning in Medical Image Analysis". In: *Medical Image Analysis* 42 (2017), pp. 60–88.

[112] G. Litjens, T. Kooi, B. E. Bejnordi, et al. "A Survey on Deep Learning in Medical Image Analysis". In: *Medical Image Analysis* 42 (2017), pp. 60–88.

[113] G. Litjens, T. Kooi, B. E. Bejnordi, A. Setio, F. Ciompi, M. Ghafoorian, L. A. W. M. van der Laak, B. Ginneken, and C. T. Sánchez. "A Survey on Deep Learning in Medical Image Analysis". In: *Medical Image Analysis* 42 (2017), pp. 60–88.

[114] Gengbo Liu, Debasis Mitra, Ella F. Jones, Benjamin L. Franc, Spencer C. Behr, Alex Nguyen, Marjan S. Bolouri, Dorota J. Wisner, Bonnie N. Joe, Laura J. Esserman, et al. "Mask-Guided Convolutional Neural Network for Breast Tumor Prognostic Outcome Prediction on 3D DCE-MR Images". In: *Journal of Digital Imaging* 34 (2021), pp. 630–636.

[115] Z. Liu, Y. Lin, Y. Cao, et al. "Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows". In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. 2021, pp. 10012–10022.

[116] M. Lobbes, M. Smidt, K. Keymeulen, R. Girometti, C. Zuiani, R. Beets-Tan, J. Wildberger, and C. Boetes. "Malignant Lesions on Mammography: Accuracy of Two Different Computer-Aided Detection Systems". In: *Clinical Imaging* 37 (2013), pp. 283–288.

[117] M. Lobbes, M. Smidt, K. Keymeulen, R. Girometti, C. Zuiani, R. Beets-Tan, J. Wildberger, and C. Boetes. "Malignant Lesions on Mammography: Accuracy of Two Different Computer-Aided Detection Systems". In: *Clinical Imaging* 37 (2013), pp. 283–288.

[118] M. Lobbes, M. Smidt, K. Keymeulen, R. Girometti, C. Zuiani, R. Beets-Tan, J. Wildberger, and C. Boetes. "Malignant lesions on mammography: accuracy of two different computer-aided detection systems". In: *Clinical Imaging* 37 (2013), pp. 283–288.

[119] J. Long, E. Shelhamer, and T. Darrell. "Fully Convolutional Networks for Semantic Segmentation". In: *Conference on Computer Vision and Pattern Recognition*. IEEE, 2015, pp. 3431–3440.

[120] Pauline Luc, Camille Couprie, Soumith Chintala, and Jakob Verbeek. "Semantic segmentation using adversarial networks". In: *NIPS Workshop on Adversarial Training*. 2016.

[121] S. Madhukumara and N. Santhiyakumari. "Evaluation of K-Means and Fuzzy C-Means Segmentation on MR Images of Brain". In: *The Egyptian Journal of Radiology and Nuclear Medicine* 46 (2015), pp. 475–479.

[122] M. Maeda and S. Tsuda. "Reduction of Artificial Bee Colony Algorithm for Global Optimization". In: *Neurocomputing* 148 (2015), pp. 70–74.

[123] G. Maicas, G. Carneiro, and A. P. Bradley. "Globally Optimal Breast Mass Segmentation from DCE-MRI Using Deep Semantic Segmentation as Shape Prior". In: *14th International Symposium on Biomedical Imaging*. IEEE, 2017, pp. 305–309.

[124] P. Mansfield. "Multi-Planar Image Formation Using NMR Spin Echoes". In: *Journal of Physics C: Solid State Physics* 10.3 (1977), pp. L55–L58.

[125] P. Mansfield. "Multi-Planar Image Formation Using NMR Spin Echoes". In: *Journal of Physics C: Solid State Physics* 10.3 (1977), pp. L55–L58.

[126] N. Maouche and J. Oudia. "Approche à Base d'Alliances dans les Graphes pour la Réduction de la Congestion dans les VANETs". MA thesis. Département d'Informatique, Université de Béjaia, Juillet 2017.

[127] D. Marr and E. Hildreth. "Theory of Edge Detection". In: *In 4th International Congress on Image and Signal Processing - Series B. Biological Sciences* 207 (1980), pp. 187–217.

[128] *Mastectomy (Definition)*. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/mastectomy. (Accessed : 24/06/2025).

[129] D. McClymont, A. Mehnert, A. Trakic, D. Kennedy, and S. Crozier. "Fully Automatic Lesion Segmentation in Breast MRI Using Mean-Shift and Graph-Cuts on a Region Adjacency Graph". In: *Journal of Magnetic Resonance Imaging* 39 (2014), pp. 795–804.

[130] S. Miah, C. A. Banks, M. K. Adams, L. Florens, K. E. Lukong, and M. P. Washburn. "Advancement of Mass Spectrometry-Based Proteomics Technologies to Explore Triple Negative Breast Cancer". In: *Molecular BioSystems* (2017).

[131] S. Michahial and B. A. Thomas. "Applying Cuckoo Search Based Algorithm and Hybrid Based Neural Classifier for Breast Cancer Detection Using Ultrasound Images". In: *Evolutionary Intelligence* (2019), pp. 1–18.

[132] F. Milletari, N. Navab, and Seyed-Ahmad S. A. Ahmadi. "V-Net: Fully convolutional neural networks for volumetric medical image segmentation". In: *2016 Fourth International Conference on 3D Vision (3DV)*. IEEE. 2016, pp. 565–571.

[133] M. Mirza and S. Osindero. "Conditional generative adversarial nets". In: *arXiv preprint arXiv:1411.1784* (2014).

[134] H. M. Moftah, A. T. Azar, E. T. Al-Shammari, N. I. Ghali, A. E. Hassanien, and M. Shoman. "Adaptive K-Means Clustering Algorithm for MR Breast Image Segmentation". In: *Neural Computing and Applications* 24 (2014), pp. 1917–1928.

[135] K. Mori, N. Niki, Y. Kawata, H. Fujita, M. Oda, H. Kim, H. Arimura, A. Shimizu, S. Noriki, K. Inai, and H. Kimura. "Applied Technologies and Systems". In: *Computational Anatomy Based on Whole Body Imaging*. Ed. by K. Hidefumi and M. Yoshitaka. 1st. Tokyo: Springer, 2017, pp. 285–352.

[136] E. Mortensen. "Interactive Segmentation with Intelligent Scissors". In: *Graphical Models and Image Processing* 60 (1998), pp. 349–384.

[137] C. K. Mummadi, M. Rosendahl, T. Brox, and V. Natarajan. "On the Robustness of Semantic Segmentation Models to Input Perturbations". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2021, pp. 888–897.

[138] J. Nahla and M. Monica. "Application of Fuzzy Neural Network for Image Tumor Description". In: *World Academy of Science, Engineering and Technology* (2019).

[139] A. Nakib, B. Daachi, and P. Siarry. "Hybrid Differential Evolution Using Low-Discrepancy Sequences for Image Segmentation". In: *26th International Parallel and Distributed Processing Symposium Workshops and PhD Forum*. IEEE, 2012, pp. 634–640.

[140] A. Nakib, H. Oulhadj, and P. Siarry. "Image Histogram Thresholding Based on Multi-Objective Optimization". In: *Signal Processing* 87 (2007), pp. 2516–2534.

[141] N.Q. Nguyen and S.W. Lee. "Robust Boundary Segmentation in Medical Images Using a Consecutive Deep Encoder-Decoder Network". In: *IEEE Access* 7 (2019), pp. 33795–33808.

[142] C. F. Nodine and H. L. Kundel. "Using Eye Movements to Study Visual Search and to Improve Tumor Detection". In: *Radiographics: A Review Publication of the Radiological Society of North America, Inc* 7 (1987), pp. 1241–1250.

[143] C. F. Nodine and H. L. Kundel. "Using eye movements to study visual search and to improve tumor detection". In: *Radiographics : a review publication of the Radiological Society of North America, Inc* 7 (1987), pp. 1241–1250.

[144] P. Orbanz and J. M. Buhmann. "Nonparametric Bayesian Image Segmentation". In: *Journal of Magnetic Resonance Imaging* 77 (2008), pp. 25–45.

[145] https://www.who.int/. (Accessed : 24/06/2025).

[146] S. Osher and J. A. Sethian. "Fronts Propagating with Curvature-Dependent Speed: Algorithms Based on Hamilton-Jacobi Formulations". In: *Journal of Computational Physics* 79 (1988), pp. 12–49.

[147] K. Ouazine. "Alliances dans les Graphes: Propriétés et Application pour la Réduction de la Saturation et de la Congestion dans les Réseaux VANETs". Thèse en Informatique. Université de Béjaia, Août 2018.

[148] K. Ouazine, H. Slimani, H. Nacer, N. Bermad, and S. Zemmoudj. "Reducing Saturation and Congestion in VANET Networks: Alliance-Based Approach and Comparisons". In: *International Journal of Communication Systems* 33.4 (2020), e4245.

[149] K. Ouazine, H. Slimani, H. Nacer, J. Oudia, and N. Maouche. "Alliance Based Approach for Reducing Saturation and Congestion in VANETs". In: *2018 International Symposium on Programming and Systems (ISPS)*. 2018, pp. 1–6.

[150] R. Pascanu, T. Mikolovs, and Y. Bengio. "On the difficulty of training recurrent neural networks". In: *International conference on machine learning*. PMLR. 2013, pp. 1310–1318.

[151] G. Pavlidis. *Segmentation of Digital Images*. Signals and Communication Technology. Springer, 2017, pp. 213–260.

[152] B. Peng, L. Zhang, and D. Zhang. "A Survey of Graph Theoretical Approaches to Image Segmentation". In: *Pattern Recognition* 46 (2013), pp. 1020–1038.

[153] C. Peng, Y. Zhang, J. Zheng, B. Li, J. Shen, M. Li, L. Liu, B. Qiu, and D. Z. Chen. "IMIIN: An Inter-Modality Information Interaction Network for 3D Multi-Modal Breast Tumor Segmentation". In: *Computerized Medical Imaging and Graphics* 95 (2022), p. 102021.

[154] G. Piantadosi. "Pattern Recognition in Breast DCE-MRI Automatic Cancer Analysis". PhD thesis. Università Degli Studi Di Napoli Federico II, Napoli, Italy, June 2018.

[155] E. J. Pruche and K. Bellenir. *Breast Cancer Sourcebook (Health Reference Series)*. Omnigraphics, 2001.

[156] *Python (Langage de Programmation) - Wikipédia*. https://fr.wikipedia.org/wiki/Python_(langage). (Accessed : 24/06/2025).

[157] Alec Radford, Luke Metz, and Soumith Chintala. "Unsupervised representation learning with deep convolutional generative adversarial networks". In: *arXiv preprint arXiv:1511.06434* (2015).

[158] S. Rahangdale, P. Keijzer, and P. Kruit. "MBSEM Image Acquisition and Image Processing in LabView FPGA". In: *International Conference on Systems, Signals and Image Processing, IEEE*. 2016, pp. 1–4.

[159] N. S. M. Raja, S. L. Fernandes, N. Dey, S. C. Satapathy, and V. Rajinikanth. "Contrast Enhanced Medical MRI Evaluation Using Tsallis Entropy and Region Growing Segmentation". In: *Journal of Ambient Intelligence and Humanized Computing* (2018), pp. 1–12.

[160] N. K. Raman and N. Sahu. "Efficient System for Devanagari Script Segmentation". In: *2nd International Conference on Computing for Sustainable Global Development - IEEE*. 2015, pp. 722–725.

[161] N. V. Rehman, J. Wang, H. Weiyan, I. Ali, A. Akbar, M. Assam, Y. Ghadi Yasin, and A. Algarni. "Edge of Discovery: Enhancing Breast Tumor MRI Analysis with Boundary-Driven Deep Learning". In: *Biomedical Signal Processing and Control* 95 (2024), p. 106291.

[162] *RIDER Breast MRI (Dataset)*. https://wiki.cancerimagingarchive.net/display/Public/RIDER+Breast+MRI. (Accessed : 24/06/2025).

[163] L. G. Roberts. "Machine Perception of Three-Dimensional Solids". PhD thesis. Massachusetts Institute of Technology, Cambridge, USA, June 1963.

[164] O. Ronneberger, P. Fischer, and T. Brox. "U-net: Convolutional Networks for Biomedical Image Segmentation". In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer. 2015, pp. 234–241.

[165] K. R. Sail. "Outcomes Associated with Chemotherapy in Elderly Patients with Early Stage Operable Breast Cancer". PhD thesis in Health Science. The University of Texas, Texas, USA, Aug. 2010.

[166] T. Salimans, I. Goodfellow, W. Zaremba, V. Cheung, A. Radford, and X. Chen. "Improved techniques for training GANs". In: *Advances in neural information processing systems* 29 (2016).

[167] L. Sani, M. Paoli, G. Raspa, A. Vispa, N. Ghavami, G. Tiberi, A. Saracini, S. Ercolani, E. Vannini, and M. Duranti. "Microwave Apparatus for Testing Breast Integrity Based on Huygens' Principle: Clinical Validation on 16 Subjects". In: *Loughborough Antennas and Propagation Conference*. Ed. by IET Digital Library. 2017, pp. 1–5.

[168] https://www.e-cancer.fr/Patients-et-proches/Les-cancers/Cancer-du-sein/Anatomie-du-sein. (Accessed : 24/06/2025).

[169] M. Semchedine. "Contribution à la Segmentation d'Images Médicales par les Algorithmes Bio-Inspirés". PhD thesis in Computer Science. Université Ferhat Abbas de Sétif, Juillet 2018.

[170] J. Senthilnath, R. Rajendra, S. Suresh, S. Kulkarni, and J. A. Benediktsson. "Hierarchical Clustering Approaches for Flood Assessment Using Multi-Sensor Satellite Images". In: *International Journal of Image and Data Fusion* 10 (2019), pp. 28–44.

[171] K. H. Shafique and R. D. Dutton. "A Tight Bound on the Cardinalities of Maximum Alliance-Free and Minimum Alliance-Cover Sets". In: *Journal of Combinatorial Mathematics and Combinatorial Computing* 10 (2006), pp. 139–145.

[172] A. Sharifian, S. F. Sasansara, M. J. Ghadi, S. Ghavidel, L. Li, and J. Zhang. "Dynamic Performance Improvement of an Ultra-Lift Luo DC-DC Converter by Using a Type-2 Fuzzy Neural Controller". In: *Computers and Electrical Engineering* 69 (2018), pp. 171–182.

[173] J. Shi and J. Malik. "Normalized Cuts and Image Segmentation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22 (2000), pp. 888–905.

[174] T. Si, D. K. Patra, S. Mondal, and P. Mukherjee. "Breast Lesion Segmentation in DCE-MRI Using Multi-Objective Clustering with NSGA-II". In: *International Conference on Innovative Trends in Information Technology (ICITIIT)*. IEEE, 2022.

[175] smallboy-code. *Breast-cancer-dataset: A DCE-MRI dataset for breast tumor image segmentation and classification*. GitHub repository. 2022. URL: https://github.com/smallboy-code/Breast-cancer-dataset.

[176] I. Sobel and G. Feldman. "A 3x3 Isotropic Gradient Operator for Image Processing". In: *Presented as a Talk Within the Stanford Artificial Intelligence Project* (1968).

[177] https://pathology.jhu.edu/breast/staging-grade/. (Accessed : 24/06/2025).

[178] L. Sun, J. He, X. Yin, Y. Zhang, J. H. Chen, T. Kron, and M. Y. Su. "An Image Segmentation Framework for Extracting Tumors from Breast Magnetic Resonance Images". In: *Journal of Innovative Optical Health Sciences* 11 (2018), pp. 1850014–1850015.

[179] G. Szabö and T. Czàràn. "Defensive Alliances in Spatial Models of Cyclical Population Interactions". In: *Physical Reviews* 64 (2001), 11 pages.

[180] https://acsjournals.onlinelibrary.wiley.com/doi/10.3322/caac.21834. (Accessed : 24/06/2025).

[181] A. Teramoto, S. Miyajo, H. Fujita, O. Yamamuro, K. Omi, and M. Nishio. "Automated Analysis of Breast Tumour in the Breast DCE-MR Images Using Level Set Method and Selective Enhancement of Invasive Regions". In: *International Workshop on Breast Imaging*. Springer, 2016, pp. 439–445.

[182] B. Thapa, S. Thapa, S. P. Aryal, A. Sedhain, and A. Shrestha. "Performance Evaluation of CNN Under Different Types of Noise for Augmented Medical Image Classification". In: *Procedia Computer Science* 179 (2021), pp. 524–531.

[183] *TSA-Algérie: Il y a Deux Types de Cancers qui se Détachent en Algérie*. https://www.tsa-algerie.com/pr-zitouni-il-y-a-deux-types-de-cancers-qui-se-detachent-en-algerie/. (Accessed: 24/06/2025).

[184] V. Vapnik. *The Nature of Statistical Learning Theory*. 1st. New York: Springer Science and Business Media, 1995, pp. 213–260.

[185] A. Vaswani, N. Shazeer, N. Parmar, et al. "Attention is All You Need". In: *Advances in Neural Information Processing Systems (NeurIPS)*. Vol. 30. 2017, pp. 5998–6008.

[186] U. Veronesi, B. Salvadori, A. Luini, M. Greco, R. Saccozzi, M. Del Vecchio, L. Mariani, S. Zurrida, and F. Rilke. "Breast Conservation Is a Safe Method in Patients with Small Cancer of the Breast. Long-Term Results of Three Randomised Trials on 1.973 Patients". In: *European Journal of Cancer* 31 (1995), pp. 1574–1579.

[187] Hongyu Wang, Jiaqi Cao, Jun Feng, Yilin Xie, Di Yang, and Baoying Chen. "Mixed 2D and 3D Convolutional Network with Multi-Scale Context for Lesion Segmentation in Breast DCE-MRI". In: *Biomedical Signal Processing and Control* 68 (2021), p. 102607.

[188] Z. Wang, J. R. Jensen, and J. Im. "An Automatic Region-Based Image Segmentation Algorithm for Remote Sensing Applications". In: *Environmental Modelling and Software* 25 (2010), pp. 1149–1165.

[189] https://fr.wikipedia.org/wiki/Sein. (Accessed : 24/06/2025).

[190] J. Wu, Z. Wang, M. Hong, W. Ji, H. Fu, Y. Xu, M. Xu, and Y. Jin. "Medical sam adapter: Adapting segment anything model for medical image segmentation". In: *Medical image analysis* 102 (2025), p. 103547.

[191] Z. Wu and R. Leahy. "An Optimal Graph Theoretic Approach to Data Clustering: Theory and Its Application to Image Segmentation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15 (2002), pp. 1101–1113.

[192] E. Xie, W. Wang, Z. Yu, et al. "SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers". In: *Advances in Neural Information Processing Systems (NeurIPS)*. Vol. 34. 2021, pp. 12077–12090.

[193] X. Xie, S. Y. Yeo, M. Mirmehdi, I. Sazonov, and P. Nithiarasu. "Image Gradient Based Level Set Methods in 2D and 3D". In: *Deformation Models*. Ed. by H. M. Gonzàlez, A. M. Torres, and J. V. Gómez. 1st. Dordrecht: Springer, 2013, pp. 101–120.

[194] Y. Xu and E. C. Uberbacher. "2D Image Segmentation Using Minimum Spanning Trees". In: *Image and Vision Computing* 15 (1997), pp. 47–57.

[195] Z. Xu and P. K. Srimani. "Self-Stabilizing Distributed Algorithms for Graph Alliances". In: *Proceedings of the 20th International Parallel and Distributed Processing Symposium* (2006).

[196] F. Yang and T. Jiang. "Pixon-Based Image Segmentation with Markov Random Fields". In: *IEEE Transactions on Image Processing* 12.12 (2003), pp. 1552–1559.

[197] J. Yao, J. Chen, and C. Chow. "Breast Tumor Analysis in Dynamic Contrast Enhanced MRI Using Texture Features and Wavelet Transform". In: *IEEE Journal of Selected Topics in Signal Processing* 3 (2009), pp. 94–100.

[198] I. G. Yero and J. A. Rodríguez-Velázquez. "Boundary Powerful K-Alliances in Graphs". In: *Ars Combinatoria* 111 (2013), pp. 495–504.

[199] H. Young, R. Baum, U. Cremerius, K. Herholz, O. Hoekstra, A. A. Lammertsma, J. Pruim, and P. Price. "Measurement of Clinical and Subclinical Tumor Response Using [18F]-Fluorodeoxyglucose and Positron Emission Tomography: Review and 1999 EORTC Recommendations". In: *European Journal of Cancer* 35 (1999), pp. 1773–1782.

[200] N. Yu, J. Wu, S. P. Weinstein, B. Gaonkar, B. M. Keller, A. B. Ashraf, Y. Jiang, C. Davatzikos, E. F. Conant, and D. Kontos. "A Superpixel-Based Framework for Automatic Tumor Segmentation on Breast DCE-MRI". In: *Medical Imaging 2015: Computer-Aided Diagnosis*. International Society for Optics and Photonics, 2015, 94140O.

[201] Q. Yu, K. Huang, Y. Zhu, X. Chen, and W. Meng. "Preliminary results of computer-aided diagnosis for magnetic resonance imaging of solid breast lesions". In: *Breast Cancer Research and Treatment*. Springer. 2019.

[202] A. H. Yurttakal, H. Erbay, T. Ikizceli, and S. Karacavus. "Detection of Breast Cancer Via Deep Convolution Neural Networks Using MRI Images". In: *Multimedia Tools and Applications* 79.5 (2020).

[203] C. T. Zahn. "Graph-Theoretical Methods for Detecting and Describing Gestalt Clusters". In: *IEEE Transactions on Computers* 1 (1971), pp. 68–86.

[204] *National Cancer Institute (Official Website)*. https://www.cancer.gov/. (Accessed : 24/06/2025).

[205] *Anaconda Distribution*. https://www.anaconda.com/products/distribution. (Accessed : 24/06/2025).

[206] Y. Zhang, S. Chan, V. Y. Park, K. Chang, S. Mehta, M. J. Kim, F. J. Combs, P. Chang, D. Chow, R. Parajuli, et al. "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". In: *Academic Radiology* 29 (2022), S135–S144.

[207] Y. Zhang, S. Li, V. Y. Park, K. T. Chang, S. Mehta, M. J. Kim, F. J. Combs, P. Chang, D. Chow, R. Parajuli, R. S. Mehta, C. Y. Lin, S. H. Chien, J. H. Chen, and M. Y. Su. "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". In: *Academic Radiology* (2020).

[208] Y. Zhang, S. Li, Z. Xu, and J. Yuan. "Novel Approach for Region Merging and Image Segmentation for Human-Computer Interaction". In: *Optical Engineering* 42 (2003), pp. 2277–2280.

[209] S. Zhao, L. Zhang, Y. Shen, and Y. Zhu. "A CNN-Based Depth Estimation Approach with Multiscale Sub-Pixel Convolutions and a Smoothness Constraint". In: *Asian Conference on Computer Vision*. Springer, 2018, pp. 365–380.

[210] X. Zhao, Y. Liao, J. Xie, X. He, S. Zhang, G. Wang, J. Fang, H. Lu, and J. Yu. "BreastDM: A DCE-MRI Dataset for Breast Tumor Image Segmentation and Classification". In: *Computers in Biology and Medicine* 164 (2023), p. 107255.

[211] X. Zhao, Y. Liao, J. Xie, X. He, S. Zhang, G. Wang, J. Fang, H. Lu, and J. Yu. "BreastDM: A DCE-MRI dataset for breast tumor image segmentation and classification". In: *Computers in Biology and Medicine* 164 (2023), p. 107255.

[212] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang. "Random Erasing Data Augmentation". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 34. 07. 2020, pp. 13001–13008.

[213] R. G. Ziegler, R. N. Hoover, M. C. Pike, A. Hildesheim, A. M. Nomura, D. W. West, A. H. Wu-Williams, L. N. Kolonel, P. L. Horn-Ross, J. F. Rosenthal, and M. B. Hyer. "Migration Patterns and Breast Cancer Risk in Asian-American Women". In: *Journal of the National Cancer Institute* 85 (1993), pp. 1819–1827.

ABSTRACT

This manuscript explores the use of hybrid deep learning architectures for breast tumor segmentation in Dynamic Contrast-Enhanced MRI (DCE-MRI), using the public BreastDM dataset. After reviewing recent state-of-the-art methods, we identified key challenges related to segmentation accuracy and model robustness in this domain. To address these issues, we investigated three configurations: a standalone TransUNet, a cGAN with a PatchGAN discriminator, and a cGAN employing a hybrid CNN–Transformer discriminator. Experimental results show that the standalone TransUNet achieves a mean Dice score of 74%, outperforming several existing models and ranking among the best-performing approaches on this dataset. The cGAN-based variants also demonstrated promising results, highlighting their ability to produce realistic and coherent segmentations. Nevertheless, our internal comparison revealed that direct supervision in the TransUNet led to more stable and efficient learning. In summary, this work offers two main contributions: it establishes TransUNet as a strong baseline for DCE-MRI breast tumor segmentation, and it provides the first empirical study of cGAN-based approaches in this context, offering useful benchmarks and insights for future research.

Keywords: Breast DCE-MRI; Tumor region; Image segmentation; TransUNet method; Deep learning; Generative Adversarial Networks; Attention mechanism.

RÉSUMÉ

Ce mémoire explore l'utilisation d'architectures hybrides d'apprentissage profond pour la segmentation des tumeurs mammaires en IRM dynamique avec injection de produit de contraste (DCE-MRI), en s'appuyant sur la base de données publique BreastDM. Après une revue des méthodes récentes de l'état de l'art, nous avons identifié plusieurs défis majeurs liés à la précision de la segmentation et à la robustesse des modèles. Pour y répondre, trois configurations ont été étudiées : un TransUNet seul, un cGAN avec un discriminateur PatchGAN, et un cGAN intégrant un discriminateur hybride de type CNN–Transformer. Les résultats expérimentaux montrent que le TransUNet seul atteint un score de Dice moyen de 74%, surpassant plusieurs approches existantes et se plaçant parmi les meilleures performances sur ce jeu de données. Les variantes basées sur les cGANs ont également produit des segmentations de qualité, soulignant leur potentiel. Toutefois, notre comparaison interne a révélé que l'apprentissage supervisé direct avec TransUNet offrait une dynamique d'entraînement plus stable et plus efficace. En résumé, ce travail apporte deux contributions principales : il valide TransUNet comme une base solide pour la segmentation des tumeurs mammaires en IRM DCE, et il présente la première étude empirique de méthodes basées sur les cGANs dans ce contexte, fournissant ainsi des repères utiles et des perspectives pour de futurs travaux de recherche.

Mots-clés : IRM mammaire DCE; Région tumorale; Segmentation d'images; Method TransUnet; Apprentissage profond; Réseaux Antagonistes Génératifs; Mechnisme d'attention.