PERFORMANCE OF THE U-NET AND SEGNET IN THE SEGMENTATION OF BREAST ULTRASOUND LESIONS

Pedro de Oliveira Vianna

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Biomédica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Mestre em Engenharia Biomédica.

Orientador: Wagner Coelho de Albuquerque Pereira

Rio de Janeiro
Julho de 2020

PERFORMANCE OF THE U-NET AND SEGNET IN THE SEGMENTATION OF BREAST ULTRASOUND LESIONS

Pedro de Oliveira Vianna

DISSERTAÇÃO SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM CIÊNCIAS EM ENGENHARIA BIOMÉDICA.

Orientador: Wagner Coelho de Albuquerque Pereira

Aprovada por: Prof. Wagner Coelho de Albuquerque Pereira
Prof. Maurício Cagy
Prof. Ricardo Cordeiro de Farias
Prof. André Victor Alvarenga

RIO DE JANEIRO, RJ – BRASIL
JULHO DE 2020

Vianna, Pedro de Oliveira

Performance of the U-Net and SegNet in the Segmentation of Breast Ultrasound Lesions/Pedro de Oliveira Vianna. – Rio de Janeiro: UFRJ/COPPE, 2020.

XI, 48 p.: il.; 29, 7cm.

Orientador: Wagner Coelho de Albuquerque Pereira

Dissertação (mestrado) – UFRJ/COPPE/Programa de Engenharia Biomédica, 2020.

Referências Bibliográficas: p. 39 – 48.

1. Breast Ultrasound. 2. Semantic Segmentation. 3. Convolutional Neural Networks. I. Pereira, Wagner Coelho de Albuquerque. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Biomédica. III. Título.

A Nildo de Oliveira Vianna,
in memoriam

Agradecimentos

Gostaria de agradecer a todos que me apoiaram nessa caminhada. Primeiramente minha esposa Letícia (e nossa cachorrinha Kira), sempre um porto seguro em minha vida, motivo de tanta felicidades e tão cheia de carinho, sempre vibrou comigo a cada conquista, desde a classificação para o mestrado, a cada nota A recebida até a aprovação na defesa. Te amo.

Minha família que me educou e me fez ser o que eu eu sou hoje, meu pai, Nelson, meus padrinhos Nildo e Sueli, minha irmã Bia e minha sobrinha Clara. Estendo meus agradecimentos também a tios, primos, avós, e todos que contribuiram para a minha formação. Agradeço também à família de Letícia que desde 2012 me acolheu muito bem, principalmente avós, tias e tios e meu sogro.

Meus amigos de sempre, Eric, Filipe, Luan e Matheus, há anos me aturando e presentes nos melhores momentos da minha vida. Meus amigos que fiz no mestrado, Bianca, Ecard, Felipe, Flávia, Gabriela, Gustavo, Lucas, Marcus Vinícius e Vinícios, entre tantos outros. Minha amiga da graduação, Eduarda, por todos os almoços e conversas sobre a vida. Meus amigos que tão sempre no Maracanã comigo, Angelo, Iago, Rodrigo, Sawlo, tantos encontros felizes tivemos em 2019. E meus amigos do colégio, fico feliz de ter contato com tantos ainda hoje, mesmo que não sejamos tão próximos.

Resumo da Dissertação apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Mestre em Ciências (M.Sc.)

DESEMPENHO DA U-NET E DA SEGNET NA SEGMENTAÇÃO DE LESÕES EM ULTRASSOM DE MAMA

Pedro de Oliveira Vianna

Julho/2020

Orientador: Wagner Coelho de Albuquerque Pereira

Programa: Engenharia Biomédica

A maior parte das biópsias realizadas em mulheres com lesões identificadas em exames de imagem para mama revela benignidade. Esse procedimento é invasivo, e demanda tempo e recursos. Dessa maneira, há interesse na diminuição dessas cirurgias desnecessárias. Para alcançar esse objetivo, o uso de diagnóstico assistido por computador como uma segunda leitura tem sido cada vez mais recorrente. Pensando nisso, apresenta-se, neste trabalho, uma implementação das redes neurais convolutivas U-Net e Segnet para segmentação de lesões em imagens de ultrassom de mama. Definir qual arquitetura é mais apropriada para essa função pode ajudar na extração de características e posterior classificação de lesões, diminuindo o número de biópsias realizadas. Foi usado um banco de imagens, obtido em parceria com o Instituto Nacional de Câncer, com 2054 imagens. Foram comparadas a segmentação automática realizada pelas redes com uma segmentação manual feita por um especialista. Dentre as duas arquiteturas propostas, a U-Net obteve melhores resultados nessa tarefa, obtendo um coeficiente dice de 86.3%, e teve tempo de treino 68.3% menor.

Abstract of Dissertation presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Master of Science (M.Sc.)

PERFORMANCE OF THE U-NET AND SEGNET IN THE SEGMENTATION OF BREAST ULTRASOUND LESIONS

Pedro de Oliveira Vianna

July/2020

Advisor: Wagner Coelho de Albuquerque Pereira

Department: Biomedical Engineering

Most biopsies performed on women with lesions identified on breast imaging tests show benignity. This procedure is invasive, and demands time and resources, thus, there is interest in reducing these unnecessary surgeries. To achieve this goal, the use of computer-aided diagnostics as a second reader has been increasing lately. Considering this scenario, the present work describes an implementation of two convolutional neural networks, U-Net and Segnet, for lesion segmentation in breast ultrasound images. Defining which architecture is most appropriate for this function can help the feature extraction and subsequent classification of lesions, reducing the number of biopsies performed. We used a dataset obtained in partnership with the National Cancer Institute, with 2054 images and compared the automatic segmentation performed by the networks with a manual segmentation made by a specialist. Among the two proposed architectures, U-Net obtained better results in this task, obtaining a dice coefficient of 86.3%, and took 68.3% less training time.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Cancer is a leading cause of death worldwide, accounting for an estimated 9.6 million deaths in 2018 according to the World Health Organization. In this scenario, breast cancer appears as one of the most common types with 2.09-millions new cases and over 0.6 million deaths [1]. In Brazil, the National Cancer Institute (INCa) estimates 62.280 new cases of breast cancer for each year of the 2020-22 triennium [2].

Detection of cancer at late stages may not provide time to effective treatment, thus early diagnosis is crucial to improve the survival rate of patients. The best strategy to early detection of cancer is screening. Currently, X-rays mammography is the most used modality for breast cancer[3].

The mammogram analysis depends on the training and experience of the radiologist, thus there is an intrinsic subjective factor in the interpretation of lesions. Size and morphological variation of the lesions may result in imprecise evaluation. Also, a high proportion of dense breast tissue hinders mammography cancer detection. For such reasons, about 10% to 30% of breast lesions are not identified on mammograms [4].

Because of inconclusive findings in mammography, breast ultrasound (BUS) has emerged as an important ally in diagnostics. The use of adjunct ultrasound is thought to be a safe and inexpensive approach to improve the sensitivity and specificity of breast cancer screening [5, 6]. BUS image analysis is done by observing morphological and texture characteristics.

In cases of suspicious abnormality or highly suspicious of malignancy, radiologists issue a clinical recommendation, which is basically to perform a biopsy. However, the likelihood of malignancy may also vary due to inter- and intraobservers image interpretation. According to the American National Breast Cancer Foundation, 80% of women who undergo breast biopsy do not have breast cancer [7]. The cost due to over-diagnosis and false positives in 40-59 year-old women is estimated as $4 billion dollars per year in the United States alone [8].

In order to mitigate the subjective factor of image interpretation and reduce the

number of unnecessary biopsies, computer-aided diagnosis (CADx) systems are used as a "second reader". Double-reading of screenings decreases the uncertainty of the specialist, increasing up to 15% the sensitivity in the detection of breast cancer [9]. Based on that, the development of the CADx systems is crucial and this can be accomplished by refinement of its steps.

Commonly, CADx systems for BUS images involve four stages: (i) image preprocessing, (ii) lesion segmentation, (iii) feature extraction and (iv) classification. The segmentation is often considered the most critical step in the system, as the subsequent operations depend on an adequate separation of the lesions from the background and other structures. However, BUS segmentation is a difficult task, owing to the variance in lesion shapes, speckle artifact, low contrast and blurry boundaries [10]. Considering the malignant tumours present more irregular boundaries, segmentation accuracy is crucial.

Automatic BUS segmentation has been extensively studied in the last two decades; either fully or semi-automatic ones. It avoids the human manual task of segmenting, saving significant time. Figure 1.1 shows the distribution of techniques employed in previous works [11].

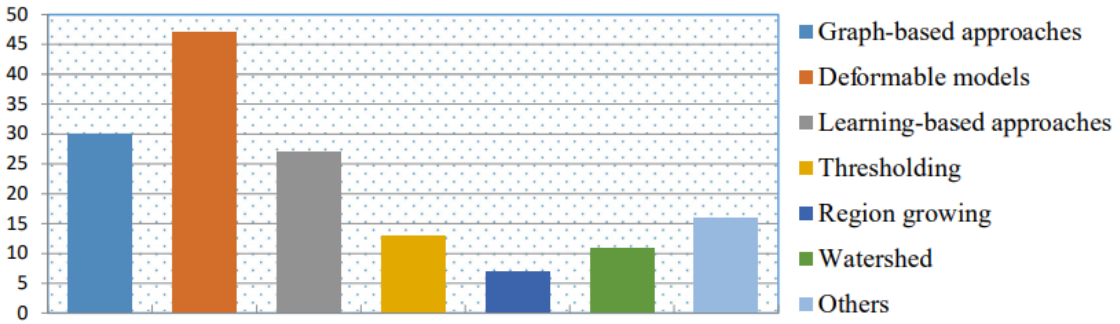

Figure 1.1: Distribution of automatic BUS image segmentation approaches; data comes from Google Scholar.

Convolutional Neural Networks (CNN) has been widely used for classification tasks in researches over the years, including in the biomedical field. Unlike traditional techniques, deep learning methods are able to learn the main features presented in an image by itself, disregarding handmade features extractors designed by a specialist in the field[12]. Using CNN to perform the segmentation procedure could help with feature extraction and posterior classification of lesions in a CADx system.

Determining which CNN accomplishes better results for segmentation tasks in BUS images should help the evolution of this field of study. For that matter, in the present work, we trained two well-known Convolutional Neural Networks, the U-Net and the SegNet, to segment images of our dataset of breast ultrasound images.

U-Net was developed at the University of Freiburg, specifically designed to work with segmentation in the biomedical field. It was awarded with two prizes in cell segmentation and was supposed to be applied easily to many more tasks. One year later, SegNet was presented as a novel architecture for general-purpose segmentation, achieving high scores while using less memory. Since then, different studies using those networks in ultrasound images were published[13–20].

The main objective of this work was to verify which CNN achieved the best segmentation results. In order to accomplish this task, we trained both architectures for segmentation of tumours on our dataset and compared the tests results with a manually delineated contour of breast lesions.

Chapter 2

Literature Review

2.1 Breast Cancer

Cancer is the name given for a group of diseases associated to abnormal cells growth, invading adjacent tissues and potentially spreading to distant organs in the body. It can affect almost any part of the body and has many anatomic and molecular subtypes that each requires specific management strategies [21].

It is the second leading cause of death globally and is estimated to account for 9.6 million deaths in 2018. Figure 2.1 (adapted from [1]) shows the estimated numbers for new cases of cancer in 2018 with lung and breast at the top. Lung, prostate, colorectal, stomach and liver cancer are the most common types of cancer in men, while breast, colorectal, lung, cervix and thyroid cancer are the most common among women [1]. The perspective in Brazil is similar, with breast cancer presenting an estimated number of 66.280 new cases for each year of the 2020-2022 triennium [2].

Cancer is a genetic disease that occurs when the information in cellular DNA becomes corrupted, leading to abnormal patterns of gene expression. The main mechanism is through the accumulation of mutations, although there is increasing recognition of the role of non-mutational changes in the process [22].

These gene mutations may be inherited, develop over time as we get older and genes wear out, or develop if we are exposed to agents that damages our genes, such as cigarette smoke, alcohol or ultraviolet radiation from the sun [23]. Once the cells start growing with damaged genes, they are more likely to pick up further mutations and less likely to be able to repair themselves [24].

Cancer cells are different from normal cells because they divide faster and out of control because they avoid the immune system, ignoring signals that inhibit proliferation or induces apoptosis. They do not develop into mature cells with specific functions, and can spread to other parts of the body through the blood or lymphatic system , as illustrated in Figure 2.2 (adapted from [23]). Similarly to normal

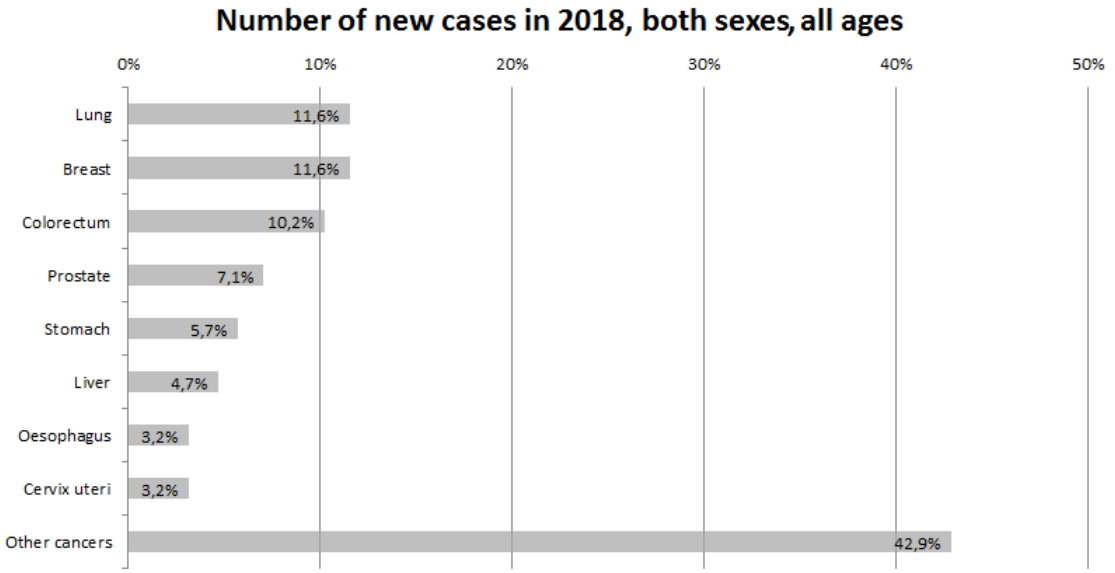

Figure 2.1: Estimated numbers of cancer new cases in 2018. Data from GLOBOCAN 2018.

cells, they need blood supply to bring them oxygen and nutrients. But as a tumour grows faster, it needs even more blood supply. So cancer cells release exceptionally large amount of vascular endothelial growth factors to form new blood vessels. This accelerated angiogenesis supports the high rate of the tumour growth [23].

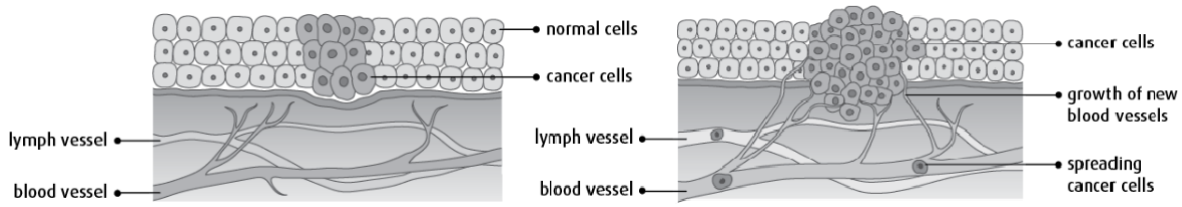

Figure 2.2: Angiogenesis and spreading cancer cells.

Most breast cancers begin in the lactiferous ducts (ductal cancers) but some start in the mammary glands (lobular cancers). The normal breast structure can be seen in Figure 2.3 [25]. Less commonly, breast cancer can begin in the stromal tissues, which include the fatty and fibrous connective tissues of the breast. Over time, cancer cells can invade nearby healthy breast tissue and make their way into the underarm lymph nodes, gaining access into other parts of the body [26]. Localized cancer leads to a 5-year survival rate of 97.5%, whereas cancer that has spread to distant organs has a 5-year survival rate of only 20.4% [27].

Although many types of breast cancer can cause a lump in the breast, not all do. Many breast cancers are found on screening mammograms which can detect cancers at an earlier stage, often before they can be felt, and before symptoms develop.

It is also important to understand that most breast lumps are benign and not

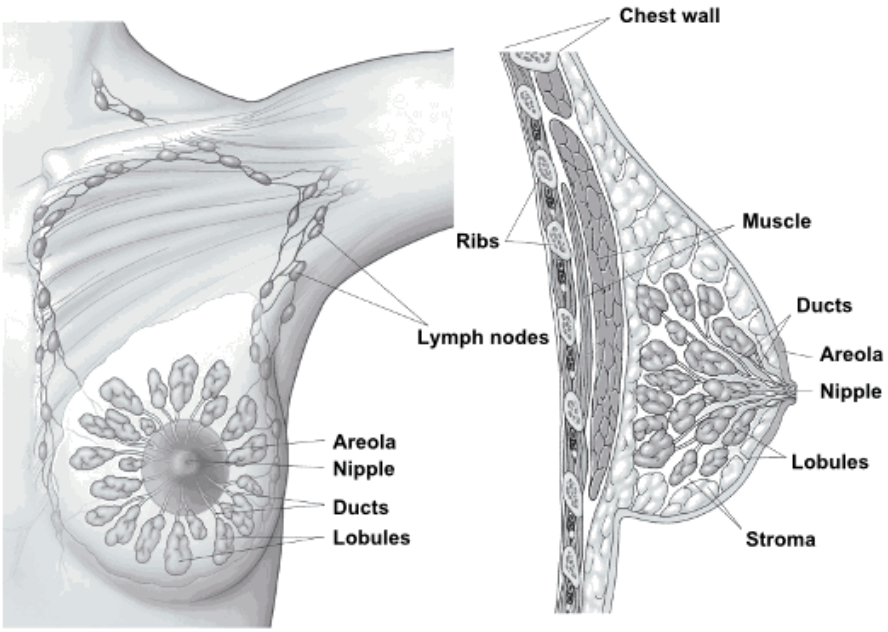

Figure 2.3: Normal breast tissue.

cancer (malignant). Non-cancerous breast tumours are abnormal growths, but they do not spread outside of the breast and they are not life threatening. But some benign breast lumps can increase a woman's risk of getting breast cancer. Detection of tumours in the earlier stages is the best way to improve the survival rate of patients. Early breast cancer usually has a very good prognosis, and the treatment may involve the combination of surgery, radiation therapy, chemotherapy, hormone therapy and/or targeted therapy.[26].

Currently, clinical exam and screening are the top-tier strategies to detect breast cancer before symptoms develop. The former is the easiest to perform, while the latter can identify non-palpable masses. Among different screening methods, mammography is the most common exam, with magnetic resonance imaging and ultrasound being used as supplemental tools [5].

Mammogram is basically an X-ray image of the breast. Differences in breast tissue generally show as different shades of grey, while microcalcifications are white. Unfortunately, the sensitivity of mammograms decreases as the breast tissue density rises. Considering that denser breasts suggest higher risks of cancer[28], this limitation of the exam is aggravated.

For that matter, other screening methods (such as the mentioned above) are used in combination to improve detection of cancer. As this work deals with breast ultrasound, the next section will focus on it.

2.2 Ultrasound

Ultrasound is the name given to sound waves whose frequency is greater than 20 kHz. These waves consist of a mechanical disturbance of molecules, creating a oscilating movement of particles in an elastic medium about an equilibrium position. The disturbance propagates through the medium at a speed which depends on the compressibility and density of the medium.

Most ultrasound applications involve transmitting short bursts, or pulses, of waves into the body and receiving echoes from tissue interfaces and scattering from inside the organs. The time between transmitting a pulse and receiving an echo is used to determine the depth of the interface. As the energy propagates, it is attenuated, scattered, and reflected, producing echoes from various interfaces[29].

Attenuation is the reduction of ultrasound amplitude during its passage through medium. This loss occurs by absorption, scatter, beam divergence, reflection and refraction. Absorption is the conversion of the wave energy into heat, due to relaxation mechanisms of particles movement.

Ultrasound applied to medicine is commonly composed of longitudinal waves, which means the direction of the propagation is the same as the disturbance. As the particles oscillate they cause a sequence of increase and decrease in local pressure. This phenomenon can be seen in Figure 2.4[29].

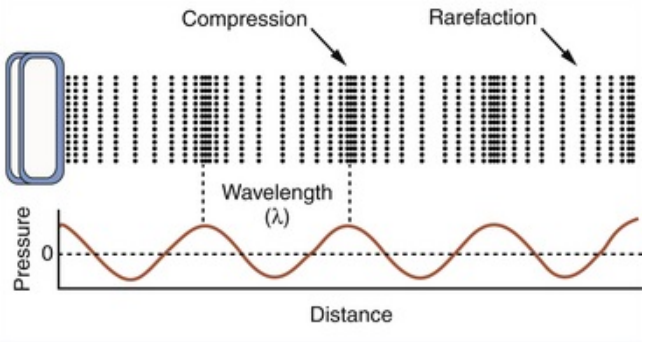

Figure 2.4: Variation of local pressure caused by longitudinal waves.

Propagation is composed of cycles of compression and rarefaction. As the pressure decreases, this energy is converted back from potential into kinetic energy. In cases of high frequency there is insufficient time for energy to change states completely, causing loss of energy in the form of heat [30]. Biological tissues are usually composed by complex structures, with distinct attenuation coefficients throughout the medium.

D'Astous and Foster[31] studied the attenuation of ultrasound on breast tissue. The results, displayed in Figure 2.5, showed that the attenuation coefficient of homogeneous regions of infiltrating duct carcinoma was higher than that of fat but

lower than that of fibrous and parenchymal tissues. They also ratify the frequency dependence of the attenuation coefficient.

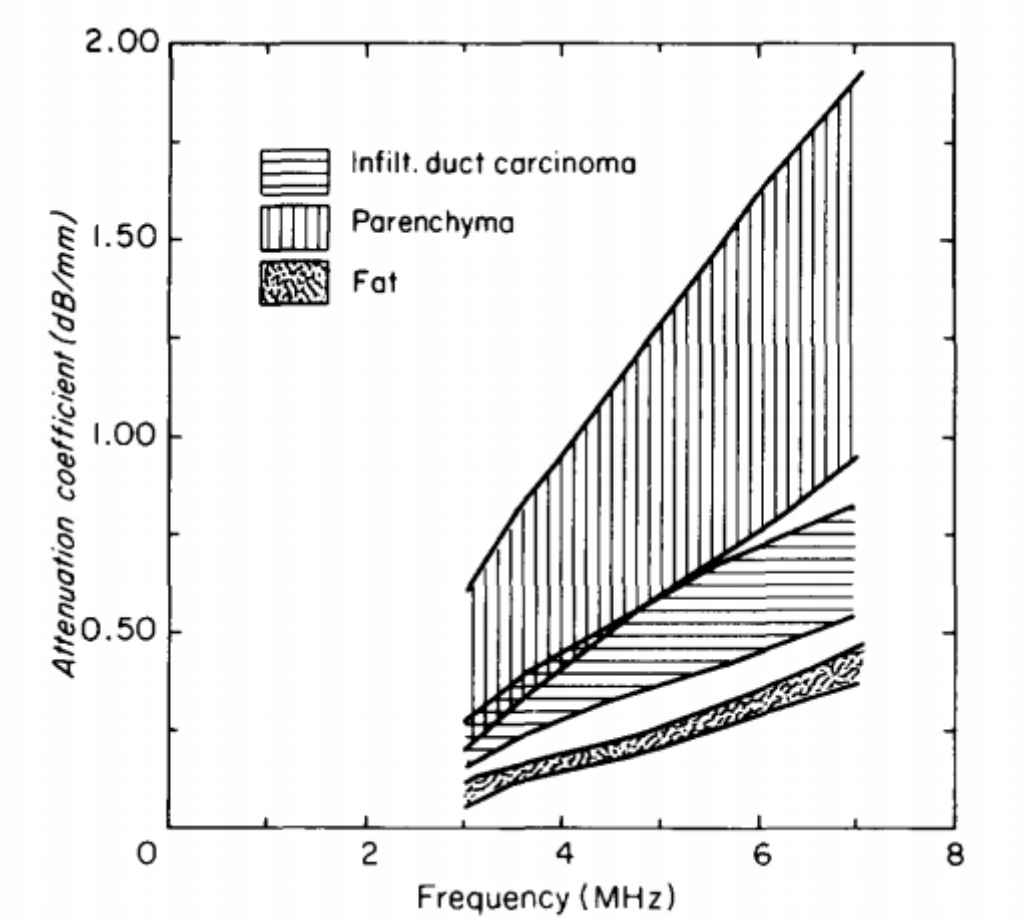

Figure 2.5: Summary of the frequency dependent attenuation coefficient measured for samples of breast fat, parenchyma and structural tissues, and infiltrating duct carcinoma.

Scattering is another phenomenon that occurs in ultrasound, acting in two manners: attenuating the main beam and producing artifacts in the experiment. It happens when the beam is incident on a rough surface or small-sized particles. This interaction is proportional to the frequency, and causes the beam to scatter in different directions.

Reflection and refraction are other phenomena inherent to the nature of wave propagation. When ultrasound is incident on a smooth interface between two different tissues, some part is reflected and some part is transmitted. The ratio of reflected intensity and transmitted intensity depends upon the characteristics of both tissues - specifically the acoustic impedance.

If the beam reaches the interface at a non-normal angle of incidence the transmitted beam will deflect obeying the Snell's Law, which is called refraction. It states that the relationship between the incident angle (θ_1) and refraction (θ_2) is proportional to the ratio of velocities v or the refractive indexes n. A scheme of reflection

and refraction of ultrasound waves is available in Figure 2.6[32].

$$\frac{sin(\theta_1)}{sin(\theta_2)} = \frac{v_2}{v_1} = \frac{n_1}{n_2} \tag{2.1}$$

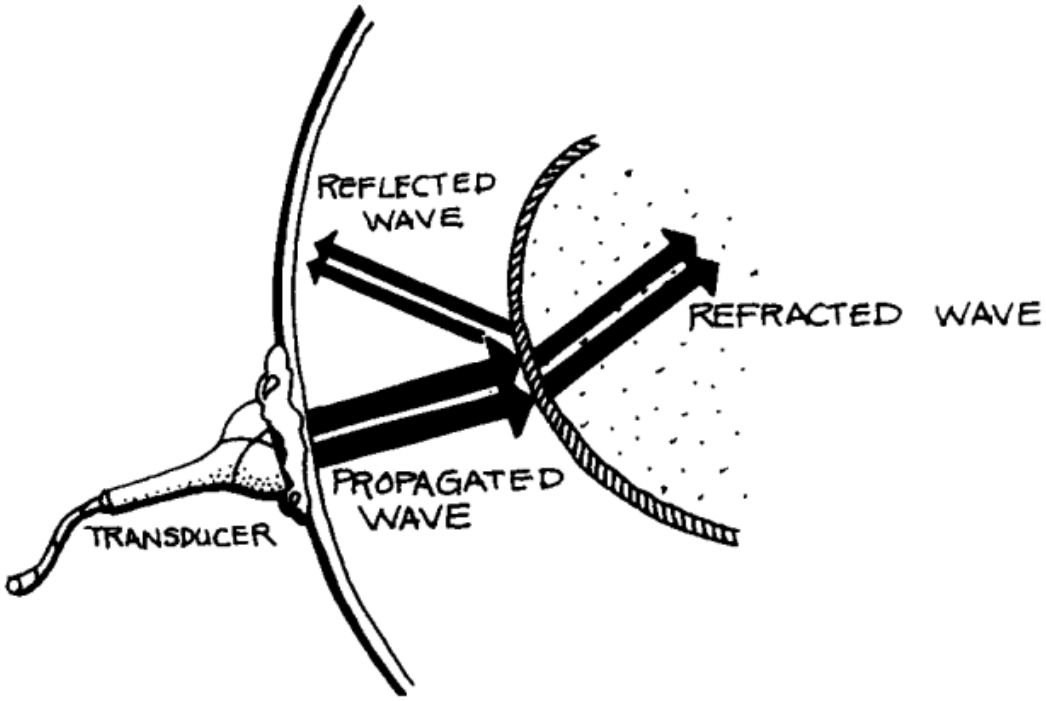

Figure 2.6: Reflection and refraction phenomena in biological tissue.

In ultrasonography, the pulse traverses through the medium and suffers these effects until an interface is encountered. The strength of the echo returning to the transducer will be dependent upon the acoustical impedance mismatch between the two tissues and the orientation of the boundary. The returning echo pulse arrives at the transducer and - considering the velocity of sound known and fixed - the position of structures can be displayed [32].

Ultrasound imaging has seen continuous development and growth over several decades. Since its introduction in the 1960s, ultrasound has found widespread application in anatomical-like imaging, blood-flow measurement, and evaluation of physiology in almost all aspects of medicine [33]. Ultrasound imaging technique has replaced or complemented a large number of radiographic and nuclear medicine procedures and has opened new areas of diagnostic investigation. As ultrasound instruments have become smaller, less expensive, and easier to use, diagnostic ultrasound has become increasingly popular among a wide variety of physicians (Table 2.1 lists many applications) [34].

Table 2.1: Examples of common uses of diagnostic ultrasound grouped by specialty.

Specialty	Applications
Radiology	Abdominal organs, peritoneum and retroperitoneum
	All paediatric applications
	Breast
	Chest and mediastinum
	Female pelvis
	Heart
	Interventional procedures (biopsy, aspiration, etc.)
	Intraoperative applications
	Neck, thyroid and parathyroid
	Obstetric examinations
	Peripheral arteries and veins
	Scrotal contents
	Soft tissues, bone, muscles, tendons and joints
	Visceral arteries and veins
Cardiology	Heart
	Large Vessels
Obstetrics	Fetus
	Placenta
	Uterus
Gynaecology	Adnexa
	Ovaries
	Uterus
Neurology and neurosurgery	Brain
	Extracranial arteries
	Intracranial arteries
Gastroenterology	Biliary system
	Gastrointestinal tract
	Liver
	Pancreas
	Spleen
Urology	Adrenals
	Kidneys
	Prostate
	Scrotal contents
	Seminal vesicles
	Ureters
	Urinary bladder

2.2.1 Breast Ultrasound

In the midst of all ultrasound imaging possibilities, breast ultrasound emerges as an important adjunct to mammography and clinical examination in the further assessment of both palpable and impalpable breast abnormalities. Recent technological advances have stimulated a resurgence of interest in the use of ultrasound as a primary screening tool, particularly in younger women in whom the theoretical radiation risks of mammography are most pertinent[35] or women with dense breast tissue as the sensitivity of mammography is low (35% - 48%) in these cases [36].

The technique used for screening must have high sensitivity for small breast, high specificity and be cost-effective, safe and acceptable by the patients. BUS satisfies all these criteria, however it is also an imperfect test, associated to a lower specificity when compared to mammography and typically requires a highly experienced operator to perform the handheld exam[37].

A common use of BUS is to distinguish if the abnormality is solid (such as a benign fibroadenoma or cancer) or fluid-filled (such as a benign cyst). However, ultrasound can also assist in differentiating among different solid masses. Various sonographic features, including a lesion margin, its shape, and its internal echo texture are evaluated[38]. Malignant tumours often infiltrate the surrounding tissue, resulting on poorly defined margins on ultrasound images, irregular and sometimes spiculated boundaries, microlobulations, hypoechogenicity, and shadowing. Benign masses, however, often have well-defined and circumscribed boundaries, and round or oval shapes with gentle bi- or trilobulations[39]. The distinction (which can be observed in Figure 2.7[40]) helps reduce the number of unnecessary biopsies, which are invasive, time-consuming and resource-expensive. However, the interpretation of images is subjected to the experience of a radiologist or physician.

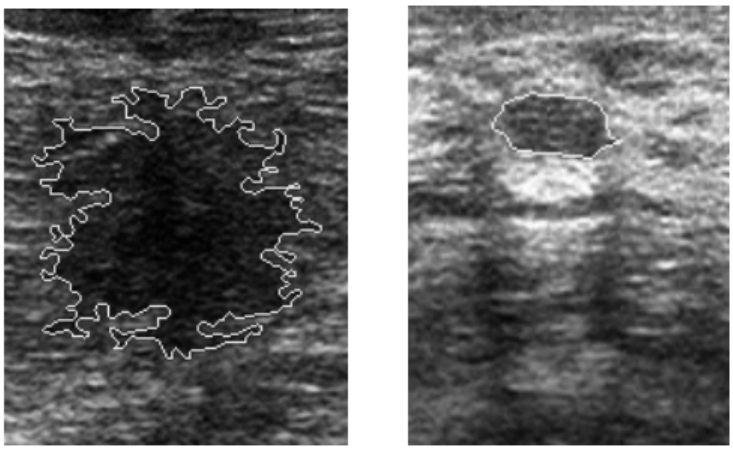

Figure 2.7: Contour established for a malign tumour (left) and a benign one (right).

Use of a second radiology reader in breast cancer screening can increase the

number of cancers detected [41], however it may be impractical due to necessity of manpower. In order to provide a cost-effective second reading of breast screening images, Computer-Aided Diagnosis (CADx) tools were developed.

2.3 Computer-Aided Diagnosis

The idea of Computer-Aided Diagnosis tool for breast sonography is to convert the visually extractable sonographic features into mathematical models and to characterize the lesions with the mathematical features based on the classification schemes[42]. Generally, ultrasound CADx systems for breast cancer involve four stages, as shown in the Figure 2.8 [43].

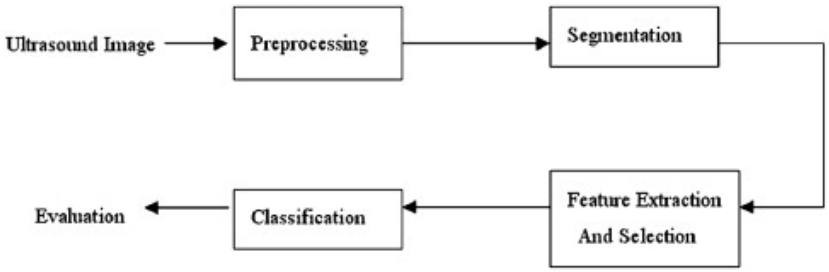

Figure 2.8: CADx system for breast cancer detection and classification.

(i)Image preprocessing:

Aims to improve the original BUS image. The image preprocessing step is usually performed with contrast enhancement followed by speckle reduction. Some enhancement techniques used are histogram equalization[44], contrast limited adaptive histogram equalization[45], sticks technique[46], fuzzy logic-based[47, 48], sigmoidal based adaptive contrast enhancement[49]. However, some systems do not consider contrast enhancement[50].

For speckle reduction the methods are subdivided in filtering[51], wavelets[52] or compounding approaches[53]. Although the former is simple and fast, it's also dependent on filter window. Working in the wavelet domain is easier for identify and properly remove noise, however it increases the operation time. Compounding approaches are simple but require hardware support[43].

Some techniques, such as nonlinear diffusion[54] enhance the image and remove speckle at the same time. It is an adaptive filter that not only preserves edges but also enhances them by inhibiting diffusion across edges and allowing diffusion on either side of the edges. Figure 2.9[55] shows an example of the importance of preprocessing stage on subsequent operations.

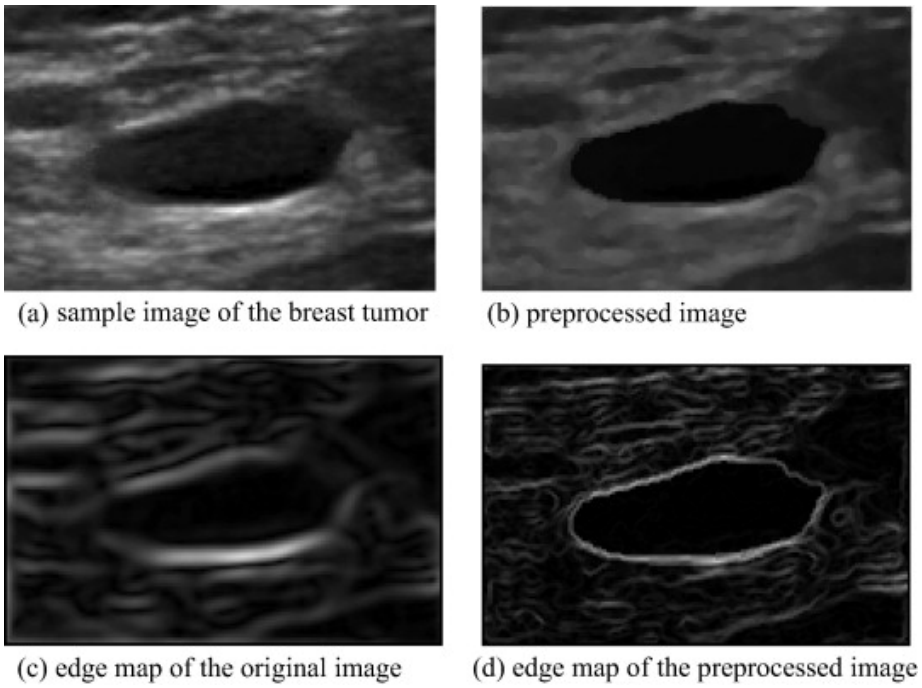

Figure 2.9: Influence of speckle reduction and edge enhancement on later edge map. Image (a) shows the BUS image and in image (b) the image was preprocessed with hierarchical fuzzy c-mean clustering. Image (c) and (d) show, respectively, the edge maps obtained from images (a) and (b), after a canny edge detection.

(ii)Image segmentation:

Divides the image into non-overlapping regions, and it will separate the objects from the background. The regions of interest (ROIs) will be allocated for feature extraction. Automatic segmentation can save much of the time required to sketch a precise contour with very high stability[56]. The goal is to locate the suspicious areas to assist radiologists in diagnoses[43]. Image segmentation is a critical and essential component and is one of the most difficult tasks in image processing and pattern recognition, and is crucial to the quality of the final analysis.

Common techniques for these tasks are histogram thresholding method[50], active contour model[57], Markov random field[58] and neural networks[59, 60]. Neural network based approaches are of increasing popularity within image segmentation, the main advantage is that it extracts the contour of the tumour automatically. However, it depends highly on the training of the network. Figure 2.10 [60] exemplifies segmentation and reinforce the influence of preprocessing in the same manner as Figure 2.9.

(iii)Feature extraction and selection:

This step is to find a feature set of breast cancer lesions that must allow the next step to accurately distinguish lesion/non-lesion or benign/malignant. The feature space could be very large and complex, so extracting and selecting the most effective

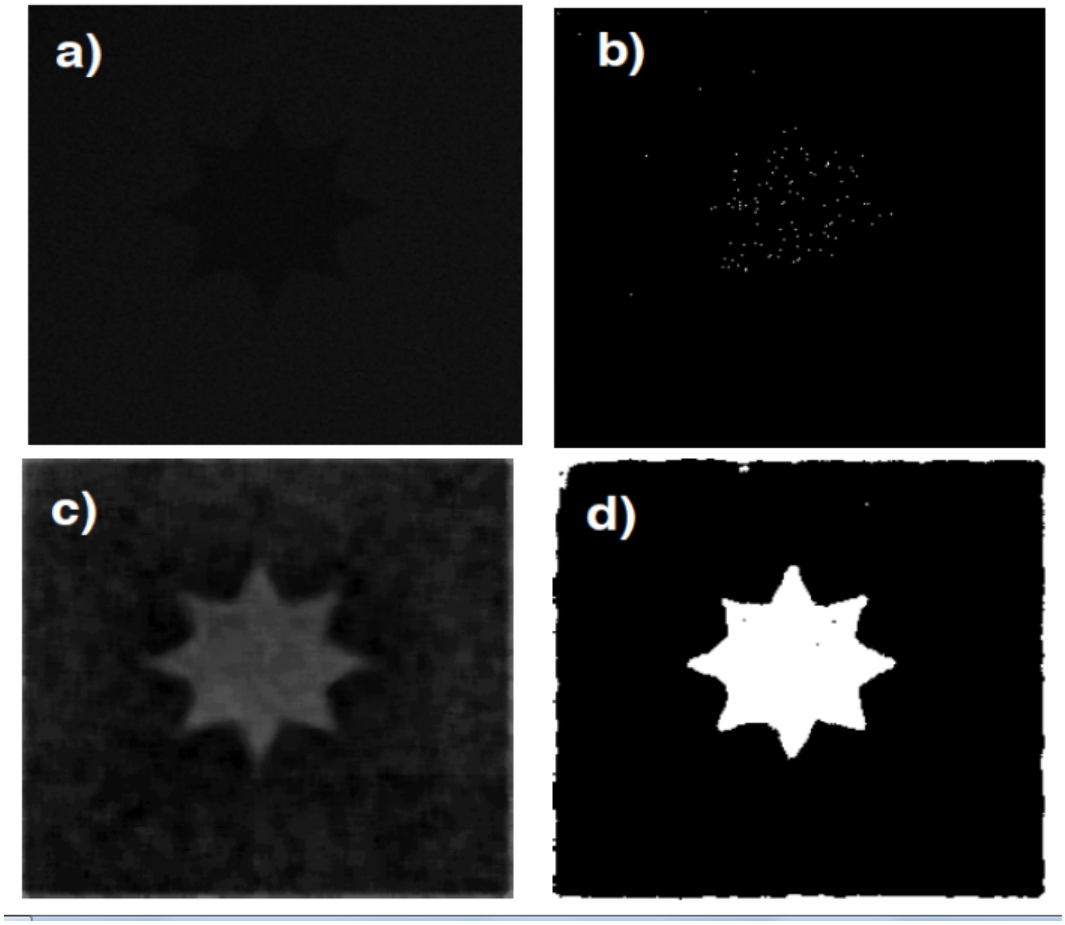

Figure 2.10: Example of segmentation using Pulse-Coupled Neural Network of a low contrast image with preprocessing (c to d) and without (a to b).

features is very important. Many authors have studied morphometric and texture features selection[61–65]. The techniques used for feature extraction are highly dependent of the previous stages.

(iv)Classification:

Based on the selected features, the suspicious regions will be classified as lesion/non-lesion or benign/malignant by various classification methods. The commonly used classifiers are linear discriminant analysis[66], neural networks[67], decision tree[68], support vector machine[69], template matching[70], and so forth.

CADx is a diagnostic aid that takes into account equally the role of the physician and the benefits of computer systems. A BUS CADx system can be an efficient computerized model and can avoid interobserver variation. [56].

Computer analysis of ultrasound images of breast lesions has been shown to improve the diagnostic accuracy of radiologists in the task of distinguishing between malignant and benign breast lesions and in recommending cases for biopsy [71–74].

2.4 Semantic Segmentation and Deep Learning

Automation of segmentation could greatly improve a CADx system if a high accuracy is achieved. As mentioned in the previous sections, the feature extraction step depends on morphometric and texture characteristics of the lesion, such as contour shape and spiculations. Thus, a correct diagnostics is more accessible with greater quality segmentation.

Many techniques have been developed for BUS segmentation. They are categorized into histogram thresholding, region growing, model-based (active contour, level set, Markov random field), machine learning, and watershed methods. All these approaches are displayed in Table 2.2 with short descriptions, main advantages and limitations (adapted from [75]).

The segmentation task can also be separated in semantic segmentation and instance segmentation. Semantic segmentation does not distinguish between elements of the same class. Instance segmentation, however, will segment and paint different objects of the same class with different colours because essentially they are different instances[12]. A comparison between both cases is presented in Figure 2.11[76]. In this work, as we are using regions of interest previously selected with a single breast tumour, we use semantic segmentation.

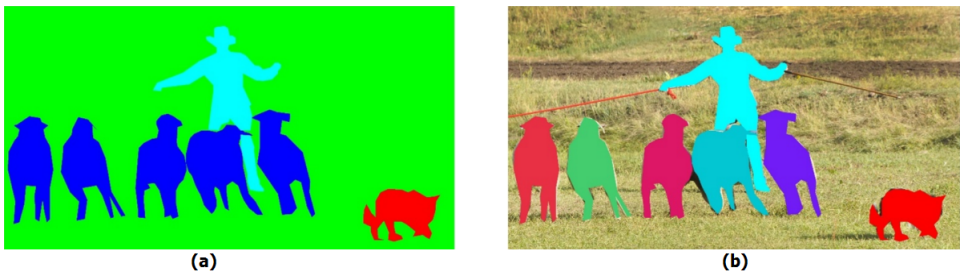

Figure 2.11: Semantic segmentation (left) and instance segmentation(right).

Semantic segmentation is currently one of the most studied topics in computer vision, with applications in medical diagnostic imaging, factory automation, remote sensing, forensics, autonomous vehicle and robot guidance[77]. The idea of this process is to recognize an image on pixel level, assigning each element to a corresponding class. This could give an explicit and meaningful description of objects within an image. The main advantages are that it is quick, consistent and automated, reducing human involvement.

Semantic segmentation algorithms are often formulated based on neural networks, encouraged by the success in classification problems. In recent years, deep learning approaches have outperformed the state of the art in segmentation tasks[78].

Deep-learning methods are representation-learning methods with multiple levels of representation, obtained by composing simple but non-linear modules in which

Table 2.2: Comparison between different approaches for segmentation of BUS images.

Methods	Descriptions	Advantages	Disadvantages
Histogram thresholding	Threshold value is selected to segment the image.	Simple and fast.	Only works for bimodal histograms and has no good results for BUS images.
Region growing	Region is grown from the seed point by adding similar neighbouring pixels.	The concept is simple. Multiple stop criteria can be chosen.	Seed point is required; sensitive to noise.
Model-based (includes active contour, level set, Markov random fields)	A model is used to formulate the lesion contour and the model is revised based on local features such as edges, intensity gradient, texture and so on.	Robust, self-adapting in search of a minimal energy state.	Time-consuming; prelabeled ROI or initial contour is required; easy to get stuck in local minima states.
Machine learning	Features to separate the lesion from the background are extracted first, and a machine learning method is trained to do the classification based on pixel-level or region-level.	Stable; different lesion characteristics can be incorporated by feature extraction.	Long training time; overtraining problem; test images should come from the same platform as the training images.
Watershed (includes marker-controlled watershed and cell-competition watershed)	Considers image as topographic surface wherein grey level of a pixel is interpreted as its altitude. Water flows along a path to finally reach a local minimum.	It ensures closed region boundaries.	Over-segmentation problem.

each module transforms the representation at one level (starting with the raw input) into a representation at a higher, slightly more abstract level. An image comes in the form of an array of pixel values, and the learned features in the first layer of representation typically represent the presence or absence of edges at particular orientations and locations in the image. The second layer typically detects motifs by spotting particular arrangements of edges, regardless of small variations in the edge positions. The third layer may assemble motifs into larger combinations that correspond to parts of familiar objects, and subsequent layers would detect objects as combinations of these parts[79].

The key aspect of deep learning is that these layers of features are not designed by human engineers: they are learned from data using a general-purpose learning procedure. Also, deep learning performs better than other machine learning algorithms when you have a large amount of data and/or that data are unstructured[79, 80].

2.4.1 Convolutional Neural Networks

Among the current deep learning techniques, the Convolutional Neural Networks (CNN's) are designed to process data that come in the form of multiple arrays, such as: 1D for signals and sequences, including language; 2D for images or audio spectrograms; and 3D for video or volumetric images[79]. Following the development and presentation of Krizhevsky's CNN for image classification (known as AlexNet)[81], researcher's interest on this technique has grown exponentially. Since then, larger and deeper networks have been proposed. Some common deep network architectures such as VGG[82], GoogLeNet[83] and ResNet[84] were used as building blocks for subsequent segmentation architectures developed.

Ronneberger, Fischer and Brox developed the U-Net[78], a CNN for biomedical applications. The architecture consists of a contracting path to capture context and a symmetric expanding path that enables precise localization. The U-Net is fast and achieves very good performance on different biomedical segmentation tasks.

Badrinarayanan, Kendall and Cipolla presented a novel CNN architecture based on the layout of the VGG-16 network, adapted to pixel-wise segmentation, named SegNet[85]. It has its structure similar to the U-Net, with contracting and expanding paths, but uses less memory. Using U-Net, Xie et al.

The basic design of a CNN consists of an arrangement of multiple layers performing operations in their inputs and forwarding the outputs to the next layer. In cases of semantic segmentation, the layers usually are: convolution, activation unit, down- and up-sampling. Also, batch normalization and dropout are optional but highly used operations. These layers are organized and repeated in different quantities depending on the intended architecture.

Parameters of convolution layer consist of a set of filters, known as kernels, to detect which features are present throughout an image. A filter is just a matrix of values, called weights, which are trained to detect specific features. The kernel moves over each part of the input and the convolution operation between the filter and that part of the image is computed. Higher values mean the feature is detected.

Figure 2.12[86] shows the convolution operation on a part of the image containing the same curve that the filter is looking for, obtaining a large number as result. Each convolutional layer has an entire set of filters and each of them produce a separate activation map. These activation maps are stacked along the depth dimension and produce the output volume.

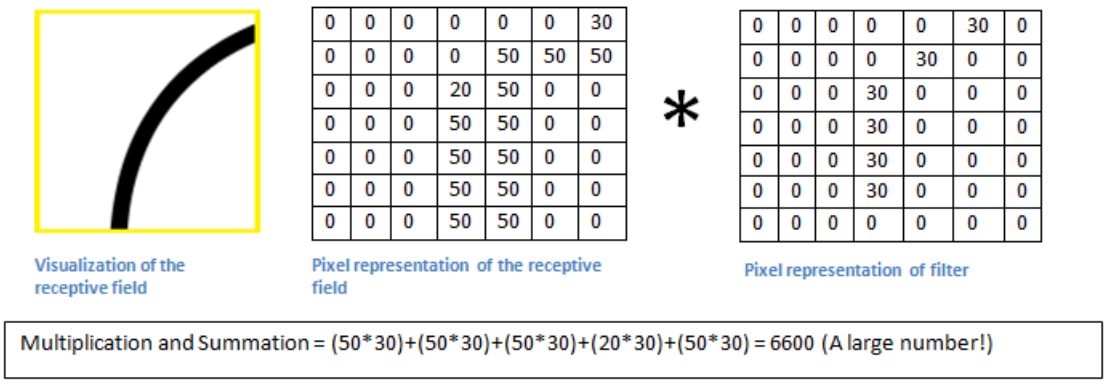

Figure 2.12: Example of convolution operation between a segment of an image and the filter checking for right-hand curve.

The output of the convolution operation between the filter and the input image is summed with a bias term and passed through a non-linear activation function. This occurs in order to introduce non-linearity to the work, as convolution is a linear operation and the input data is non-linear. The activation layer will apply an element-wise activation function, leaving the size of the volume unchanged.

Following convolution and activation comes the downsampling operation. The goal is to reduce dimensionality and to allow assumptions to be made about features contained in the sub-regions. The pooling function applies a window function to the input patch, and computes the maximum value in the neighborhood, the result is a feature map of lower resolution. Figure 2.13 [87] displays a scheme for downsampling and max pooling. This shrinkage of spatial size reduces the amount of parameters and computational time in the network.

The upsampling of the feature map is performed by an "up-convolution", also called transposed convolution, increasing the resolution of the features before the last layer (with connections to all activations in the previous layer) outputs a segmentation map.

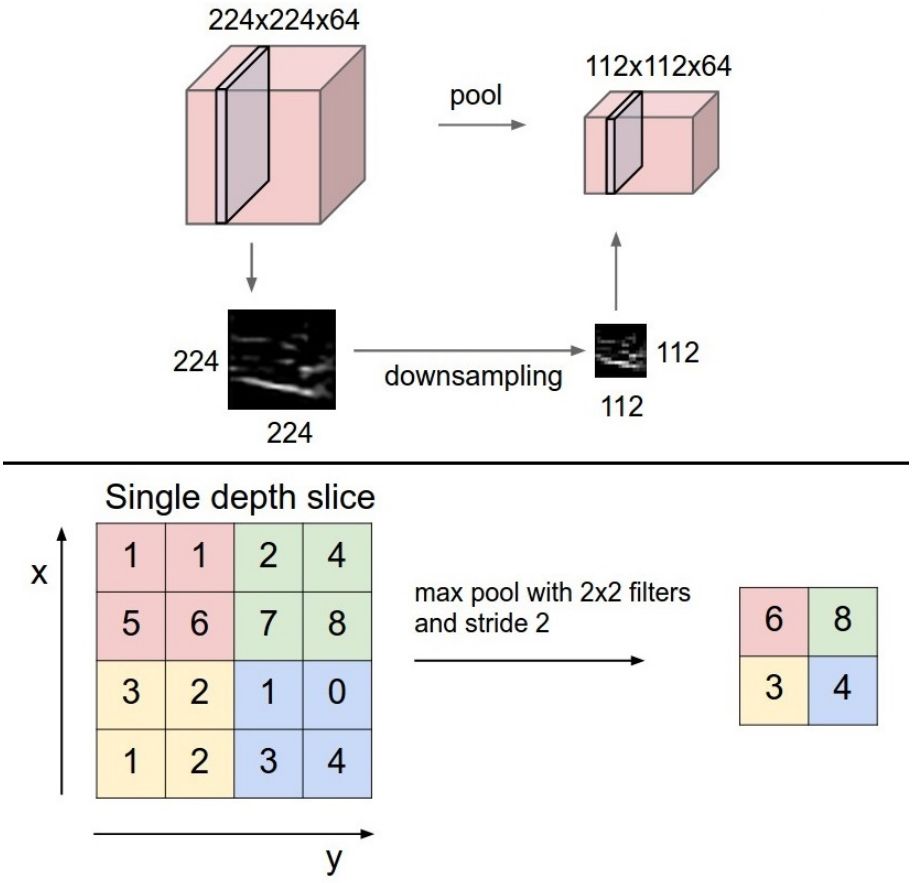

Figure 2.13: Downsample by a pooling layer (top) and example of max pool (bottom).

Dropout is a technique for improving neural networks by reducing overfitting. The key idea is to randomly drop units ("turn off" the neuron) from the neural network during training, preventing units from co-adapting too much[88]. By removing some neurons, our network becomes simpler, with less parameters, and the preserved neurons become much more specialized in the task, once they can not rely on a neighbour neuron[12].

To increase the stability of a neural network, batch normalization adapts the output of a previous activation layer by subtracting the batch mean and dividing by the batch standard deviation. It was developed in order to accelerate the training of deep networks, by adjusting and scaling the activations[89]. Similar to dropout, it can be interpreted as a way of regularizing a neural network by adding noise to its hidden units. In certain cases, the use of batch normalization makes dropout unnecessary (or at least redundant).

2.5 Related Works

As of late, deep learning has turned into a predominant research technique in various fields, and several segmentation approaches dependent on convolutional neural networks have been acquainted with medical imaging. Many authors consider U-Net, FCN-AlexNet and Mask R-CNN the state of the art architectures in image segmentation. However, there are not so many papers on breast ultrasound.

Xie et al.[13] conducted the image segmentation by modifying some structures of the aforementioned methods to adapt to their dataset of breast lesions on ultrasonography. They achieved a dice score of 0.692 using U-Net. Comparing with the different architectures of the work, the U-Net performed better in cases with irregular contours but had difficulty detecting small lesions or resolving in complex background. The best results were obtained using a modified Mask-R CNN, with a dice coefficient of 0.827.

Hu et al.[14] also studied the performance of different architectures of CNNs in breast ultrasound image segmentation. There was a dice score of 0.645 between U-Net and the ground truth labels, the lowest value of all techniques used. However it presented a standard deviation of 26%, a high value meaning that it performed much better in some cases than in others. Similar to the work mentioned before, the U-Net was great in depicting irregular shapes but failed to detect small objects. The best result (dice score = 0.890) was achieved by combining a dilated fully convolutional network with a phase-based active contour model.

Almajalid et al.[15] developed a U-net-based segmentation framework for breast tumors using BUS. With preprocessing, data augmentation, U-net training, and postprocessing steps, the group obtained a dice coefficient of 0.825.

Zhuang et al.[16] proposed a method based on the conventional U-Net, but the plain neural units are replaced with residual units to enhance the edge information. This Residual-Dilated-Attention-Gate-UNet achieved a dice score of 0.847, while the traditional U-Net scored 0.820. In addition, the SegNet was tested in the same dataset and obtained a dice score of 0.817.

Negi et al.[17] continued the work of Zhuang et al., proposing a Generative Adversarial Network model (GAN) using the RDAU-Net as the generator module, achieving a dice score of 0.884. Han et al.[18] also used a GAN for BUS lesion segmentation. In their work, they obtained a dice score of 0.871 using this architecture, while the U-Net scored 0.808.

Ghosh et al.[19] proposed a technique using U-Net and ResNet in parallel. Six convolutional blocks are coupled at the beginning as a feature extractor of the image and this set of features is fed into a combination of both architectures. This work achieved a dice score of 0.937 using the mentioned configuration, while the U-Net

obtained a dice of 0.765.

Another work in BUS segmentation that allows a comparison between U-Net and SegNet was done by Singh et al.[20]. In that work, the U-Net performed way better than the SegNet. These networks obtained dice coefficients of 0.883 and 0.504, respectively. The proposed network – a conditional generative adversarial network with an atrous convolution layer and a channel-wise weighting block model – outperforms the FCN, SegNet, UNet and other segmentation models in terms of Dice, achieving the top score of 0.938.

Other researches on Breast Ultrasound lesions segmentation using different techniques were presented in the last years. Flores and Pereira[49], from our own group, proposed a segmentation procedure using contrast enhancement and watershed transformation, acquiring a dice score of 0.895. Xian et al.[90] compared five different segmentation techniques using a dataset of 562 images. In the five approaches, two of them are graph-based approaches, one is ANN based approach, one is a level set-based segmentation approach, and one is based on cell competition. This last mentioned method is semi-automatic, and yielded their best result, with 0.880 Dice. Kriti et al.[91] proposed a non-linear iterative filter followed by active contour to segment lesions on BUS images. They obtained a dice coefficient of 0.886.

In spite of the progress, this is still an open research topic, due to challenges related to the inherent presence of noise and low contrast of images, sensitivity of current methods to the used image-acquisition method, equipment, and settings, and the lack of large open datasets of annotated images for training purposes[92].

Next section contains more detailed explanations about the techniques proposed for this work, as well as information about the BUS image database, hardware and software to be used for the project.

Chapter 3

Materials and Methods

3.1 Image Dataset

The BUS dataset consists of 2054 images from 659 female patients of ages from 16 to 89 years, with an average of 47 years. The images were acquired during routine breast diagnostic procedures at the National Cancer Institute (INCa) of Rio de Janeiro, Brazil. The INCa Research Ethics Committee has approved this study (protocol 38/2001) with prior consenting of the analyzed patients. All the cases were histopathologically proven by biopsy, where 1351 (65.8%) images presented benign lesions and 703 (34.2%) images had malignant tumours.

The image dataset was acquired using three kinds of ultrasound machines with linear array transducers: Logiq 5 (General Electric, Milwaukee, Wisconsin, USA) at 12 MHz, Logiq 7 (General Electric, Milwaukee, Wisconsin, USA) at 10 MHz, and Sonoline Sienna (Siemens, Erlangen, Baviera, Germany) at 7.5 MHz. As the beam resolution is about 0.5 mm and the pixel size is 0.33 mm/pixel, the image resolution is nearly 1.5 pixels. Also, the images were captured directly from the 8-bit video signal (i.e., 256 grey levels) and saved in TIFF format.

Every single one of the lesions had its contour manually delineated by a senior radiologist with nearly 15 years of experience in breast ultrasound interpretation. Those labels are considered the gold standard in the present work when analyzing the networks performances. An example of the image-label pair is presented on figure 3.1.

The data was randomly split between training, validation and test sets, divided in 70%, 10% and 20%, respectively. The sets are normalized with the mean (μ) and standard deviation (σ) values from all images in the training set. It is worth noting that we should use only the mean and standard deviation from the training data and apply the equations 3.1, 3.2 and 3.3 on all separate datasets. The reason is that, by machine learning principles, the test set should be reasonably representative of the

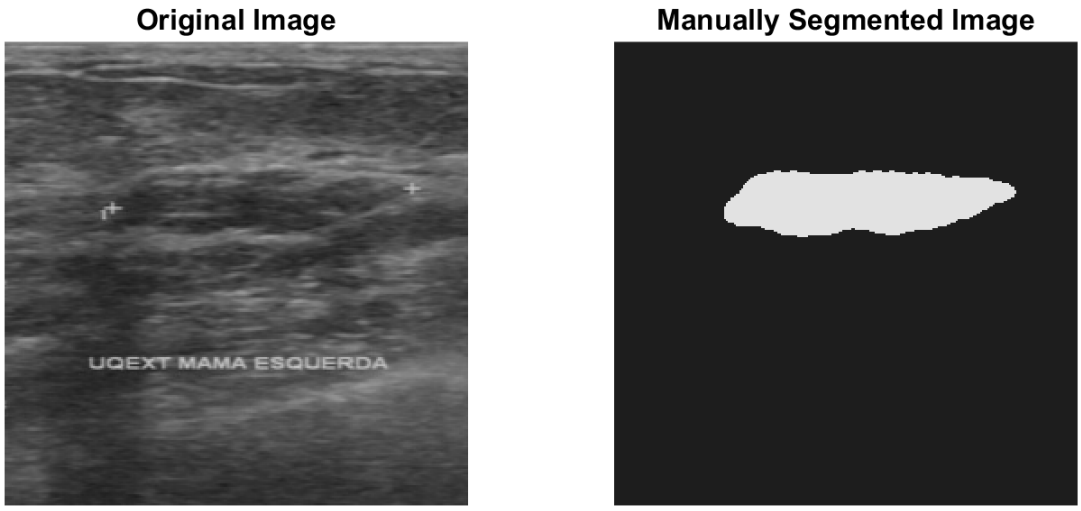

Figure 3.1: Original image from our dataset along a manually labeled lesion.

training set, consequently, we have to assume that the mean and standard deviation are the same.

$$X_{train} = \frac{X_{train} - \mu_{train}}{\sigma_{train}} \tag{3.1}$$

$$X_{validation} = \frac{X_{validation} - \mu_{train}}{\sigma_{train}} \tag{3.2}$$

$$X_{test} = \frac{X_{test} - \mu_{train}}{\sigma_{train}} \tag{3.3}$$

A small dataset - such as ours - relies heavily on data augmentation to artificially generate more data. In this work, we use two affine transformations to augment our data: translations and rotations. The operations we apply to the original images do not change the image structure. The data augmentation process is performed only on the training set, enabling a more robust and generalized model of the network. There is no procedure on the validation and test sets, as they have to represent unseen data to the model. Table 3.1 show a summary of the data augmentation method.

Table 3.1: Data augmentation table.

All Data	Training Data	Translations	Rotations	Augmentation Factor	Augmented Training Data
2054	1437	[(-1,-1),(-1,0), (0,-1),(0,0)]	[-1,0]	8	11496

Starting with the original set of 2054 images and labels, we use random 1437 images pairs as training images. For each image, we apply a translation pair (x, y) on each pixels, and apply a rotation angle a, in degrees. As we have four translation

pairs and two angles, we are augmenting the total training set size in a factor of 8, generating 10059 new images, resulting in a final training set of size 11496.

As the training is supervised, we show the labels to the network during the training in order to minimize the loss. Unlike test data, validation data is used to tune the hyper parameters. After the test stage, we compare the segmentation map of the network with the manual label to analyze the similarity between the output and the gold standard.

3.2 Hardware and Software

The testing platform considers a LINUX-based computer (64-bit Ubuntu release 16.04.6 LTS) with an Intel Core i7-7740X CPU @ 4.30GHz (Intel, Santa Clara, California, USA), 32 GB of RAM and a GeForce GTX 1080 Ti with 33 MHz clock speed (Nvidia, Santa Clara, California, USA). The Nvidia driver version was 390.59. Programming language is Python 3.6.8 with TensorFlow and Keras libraries for deep learning.

3.3 Convolutional Neural Networks

3.3.1 Functions and Parameters

Before explaining the layout of the tested networks, it is important to clarify some of the operations and functions utilized in the present work. An introductory explanation was given in section 2.4.1, and in this section we will delve into some specifics. First of all, the convolution operation must come with the choice of the filter size, stride and padding.

The filter, also known as kernel, is usually a N-by-N matrix, such as 3×3, 5×5, and so on. The stride controls how the filter convolves around the input volume. For example, the stride being set as (3,3) has the effect of moving the filter three pixels right for each horizontal movement of the filter and three pixels down for each vertical movement of the filter when creating the feature map. Lastly, padding defines how the filter deals with the borders of the input. Zero padding adds zero values around the figure, as the pixels on the edge of the input originally would only be exposed to the edge of the filter. By adding a border with zero value, the filter starts outside the frame of the image, giving the pixels on the border of the image more of an opportunity for interacting with the filter. The output's size depends on this three choices, following the formula in equation 3.4

$$O = \frac{I - K + 2P}{S} + 1 \tag{3.4}$$

Where O is the output dimension (height or length), I is the input dimension (height or length), K is the filter size, P is the padding, and S is the stride. An unpadded convolution will crop away some of the borders if the filter size is larger than 1.

A neural network without an activation function would simply work as a linear regression model, having limited power and not performing well on complex tasks. Acknowledging this, the activation function is used to convert a input signal of a node to an output signal, introducing a non-linearity to the model. We used three different activation functions in the present work.

First, the rectified linear unit (ReLU) is the most popular function in deep learning nowadays. It is a very simple and efficient unit, being easy to compute and fast to converge. f(z) is zero when z is less than zero and f(z) is equal to z when z is above or equal to zero. The downside for being zero for all negative values is the "dying ReLU", the problem when ReLU neurons become inactive and only output 0 for any input[93]. Lower learning rates often mitigates the problem. If not, other function must be used. Parametric ReLU (PReLU) is a variation of ReLU that, instead of rectifing values below zero, it uses a slope with inclination A as a parameter for the neural network to figure out itself: f(z) is equal to Az when z < 0. Xu et al. [94] compared multiple activation units - including PReLU - and obtained superior results to those with the original ReLU, using datasets with over 25 thousand images. The graphs for each function are presented in figure 3.2.

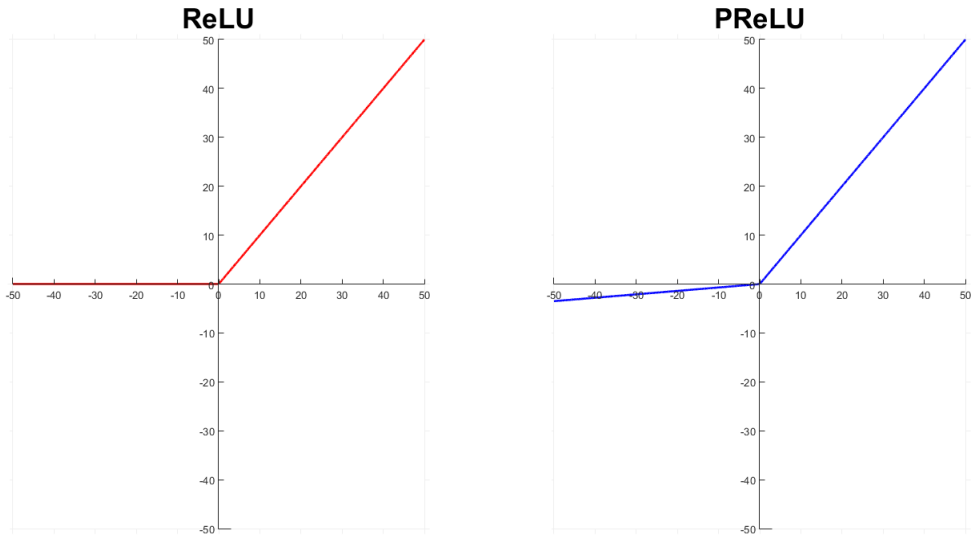

Figure 3.2: Graphs of activation functions. Left: ReLU. Right: PReLU.

The ReLU layers are meant to be used only within hidden layers of the network. In the output layer we use the third activation function of this work, the sigmoid. This function is used for binary classification, being responsible for the final operation of the network, i. e. outputting the segmentation map. It will take any real

number and return the output value in the range of 0 to 1, according to the equation 3.5.

$$f(z) = \frac{1}{1 + e^{(-z)}} \tag{3.5}$$

The downsampling is performed by the pooling layer. The most common method for that is the max pooling, used as an example previously (figure 2.13), in such a way that the maximum feature response within a given sample size is retained. The decrease in dimension is dependant on the filter size and stride used. On the other hand, the upsampling is literally a sample rate conversion opposed to downsampling. It is commonly performed by a transposed convolution or creating a bigger matrix with artificial data. The upsampling step increases the dimension of the input by the factor set in filter size and stride.

Two very important functions in the convolutional neural networks are the loss function and optimizer. The main objective of deep learning is to reduce the difference between the predicted output and the actual output. This is called the loss, and the goal is to minimize the function by finding the optimized value for the weights. The loss is summed up over all the pixels in a mini-batch and the optimizers update the weight parameters to minimize the loss function. The amount that the weights are updated during training is referred to as the step size, or the "learning rate".

In the original U-net and Segnet papers, the authors use a weighted cross-entropy [78, 85] as the loss function. However, when the level of pixel imbalance increases, Dice has better results than cross-entropy [95]. We tested both functions in the present work and compared their results in both networks. Also, in this work we used the adam optimizer. Adam is computationally efficient and requires less memory, and its characteristics makes it the most popular gradient descent optimization algorithms.

The last essential parameter to set when training the models is the dropout rate. The optimal dropout rate is allegedly 50%. A lower probability would have minimal effect and a extreme one should result in under-learning by the network.

3.3.2 Architectures

U-Net

Figure 3.3 [78] illustrates the architecture of the U-Net proposed originally. It is easy to understand origin of the name.

The input image enters the left portion, called contraction path. It consists of the repeated application of two 3×3 convolutions (unpadded convolutions), each followed by a activation unit (ReLU or PReLU) and a 2×2 max pooling operation

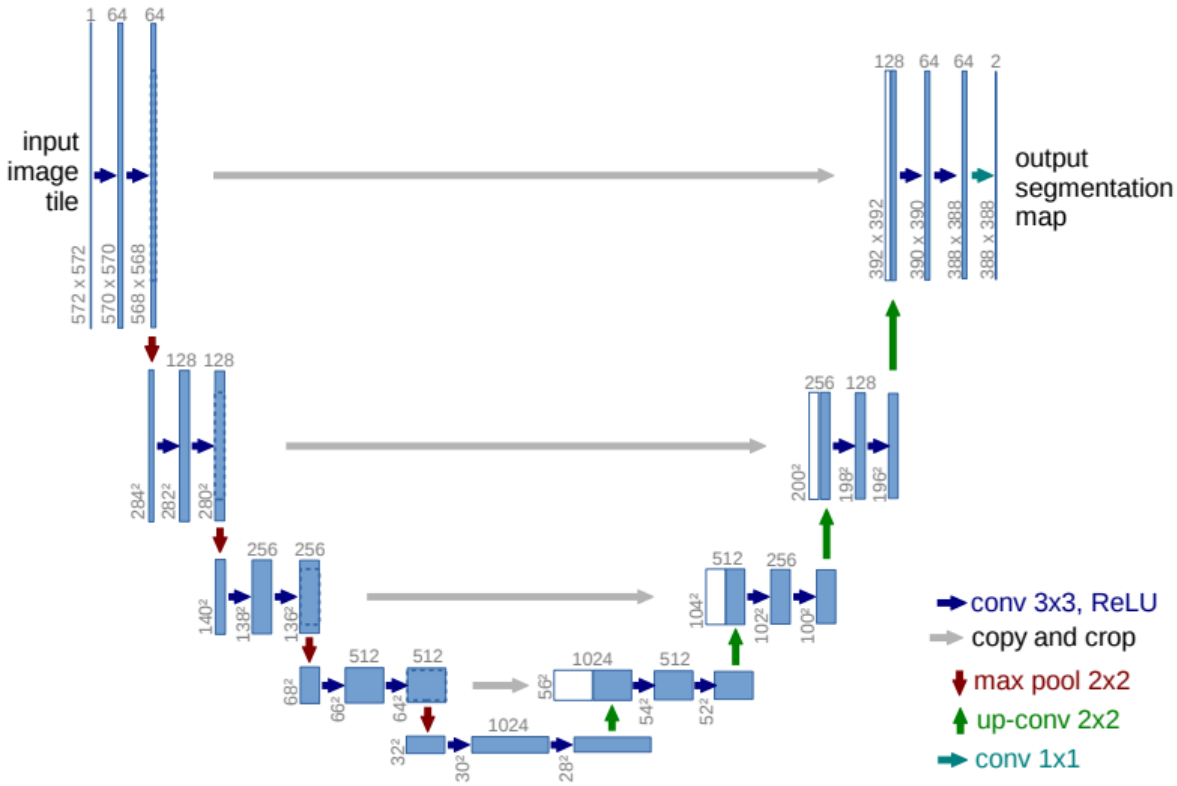

Figure 3.3: U-net architecture. The x-y-size is provided at the lower left edge of the box. White boxes represent copied feature maps. The arrows denote the different operations.

with stride 2 for downsampling. As the convolutions are unpadded, the border are cropped. In the first layer, for example, the size of the input changed from 572×572 to 570×570.

Each convolutional layer has an entire set of filters (the number of channels is denoted on top of the boxes in figure 3.3) and each of them produce a separate 2-dimensional activation map. These activation maps are stacked along the depth dimension and produce the output volume. The max pool operation with size 2×2 and stride 2 means the window do not overlap regions and the output will be downsampled by a factor of 2.

In the expansive path (right portion of the U-net) the steps are similar. It consists of an upsampling of the feature map, a concatenation with the matching feature map of the left path and two 3×3 convolutions followed by the activation. The merger with the map from the contraction map is the uniqueness of the U-Net, other networks handle it differently. The purpose of this expanding path is to enable precise localization combined with contextual information from the contracting path. This helps to predict a good segmentation map. The last layer produces an output where we have the segmentation result after a 1×1 convolution, with the desired number of classes from the 64-component feature map.

SegNet

The overall architecture of the SegNet is similar to the U-Net, as can be seen in figure 3.4. The encoder part corresponds to the first thirteen convolutional layers in the VGG-16 network[82].

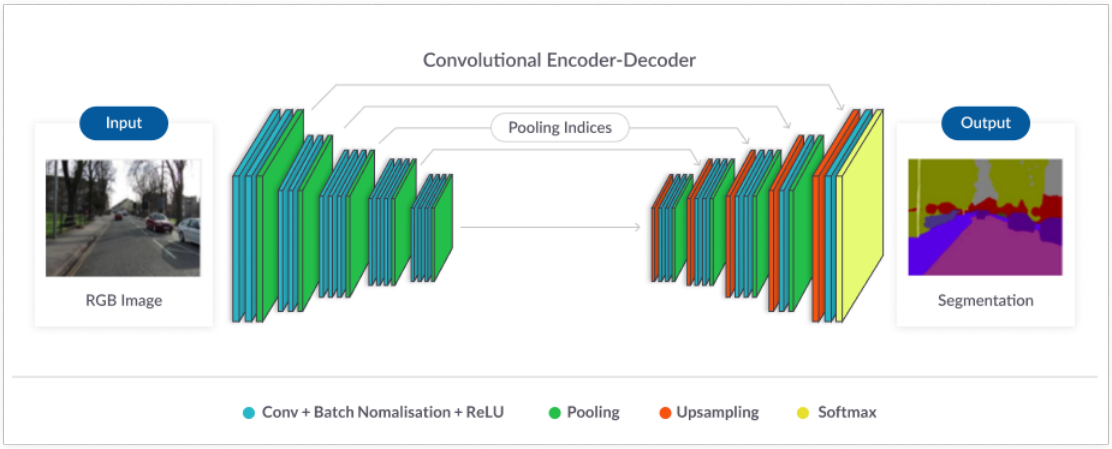

Figure 3.4: SegNet architecture. Encoding branch based on VGG-16 and decoder branch uses the pooling indices from de encoder.

The number of convolutions is different from the U-Net. There are two convolutions followed by batch normalization and activation before the pooling operation in the first two steps, and three consecutive convolutions in the following three steps. The decoder segment is mirrored in relation to the encoder, the only differences are the obvious upsampling as opposed to downsampling and the last convolution being a 1×1 for a pixel-wise classification. In the original work, the last activation function was the softmax, while we use sigmoid.

The major innovation of SegNet is the reuse of the pooling indices. The decoder upsamples its input feature maps using the memorized max-pooling indices from the corresponding encoder feature maps, instead of transferring the entire map like U-Net does. This effectively uses less memory from the system.

3.3.3 Training, Validating and Testing

In both networks the procedure is the same. We trained each network many times, using different parameters and functions when building the models. Essentially, the parameters we changed were batch size (16, 32 and 64), dropout rate (0.33 and 0.5), kernel size (3,3; 5,5 and 7,7) and learning rate (10^{-5} and 10^{-6}), and the functions were altered between ReLu and PReLU for activation and between dice and crossentropy for loss. Also, in U-Net we used a thousand epochs, while in SegNet we used two thousand. This was done because the SegNet was much slower to stabilize.

When training the network, we input 11496 images for training and 205 images for validation. While the model tries to fit the training data, the validation set is working to fine-tune the parameters. Only the model with the highest validation score is saved. In each training epoch, the training data is shuffled to disregard any bias in the presentation order. After that, we evaluate the model with the test set of 412 images.

During the testing of the model, every result image needs to be compared with the ground truth. This comparison is achieved using dice score, or dice similarity coefficient(DSC). This index accounts for the true positives(TP), false positives(FP) and false negatives(FN) between the manually labeled images and the output of the network.

$$DSC = \frac{2TP}{2TP + FP + FN} \tag{3.6}$$

Next chapter contains the description of our tests results and discussion.

Chapter 4

Results and Discussion

In this chapter we will reveal and discuss our tests results. Each test tries to achieve a model that performs better on the test set than the previous one. We tried changing some parameters as mentioned in section 3.3.3 until we obtained the best configuration for each network. Then, we compared the results of using different functions for loss and activation.

4.1 U-Net

The configuration that yields the best results in the U-Net models is presented on the table 4.1.

Table 4.1: Best parameters when compiling the U-Net models

Dropout Rate	**Batch Size**	**Filter Size**	**Learning Rate**
0.5	32	(3,3)	10^{-5}

With four possible configurations when we change the activation and the loss functions, table 4.2 resumes all four results.

Table 4.2: U-Net tests final results. The fourth column shows average dice score in the training set with standard deviation. The fifty column shows training time in hours.

#	**Activation Unit**	**Loss Function**	**Dice Score**	**Training Time (h)**
Test 1	ReLU	Dice	0.863 ± 0.141	33.06
Test 2	ReLU	Cross Entropy	0.855 ± 0.148	33.06
Test 3	PReLU	Dice	0.859 ± 0.133	45.83
Test 4	PReLU	Cross Entropy	0.851 ± 0.143	45.83

All results are similar, however some observations can be made. Most noteworthy is the time to train. While every configuration is capable of outputting 412 labels in

less than one second, the training time increases in almost 40% with PReLU activation units. No distinguishable alteration in training time occurs when comparing the loss functions.

Analyzing the dice scores of the tests, we can conclude that the biggest difference between all results is as low as 1.2%. Using a Student's t-test, the p-value between tests is lower than 5% when comparing tests 1 and 2 (p-value = 0.02) and tests 3 and 4 (p-value = 0.01). Using PReLU as opposed to ReLU does not alter the quality of the segmentation output. Meanwhile, the gap between dice and cross entropy is higher. The models with dice function segment most of the lesions in a slightly better manner than the others. Combining this with the lower training time using ReLU, the Test 1 configuration can be considered the best in the present work.

Among the 412 images on the test set, 71.1% (293) of those had dice scores above average, while 9.5% (39) were outliers, with worse labels, during Test 1. This can be seen in figure 4.1, with the dice scores for each image.

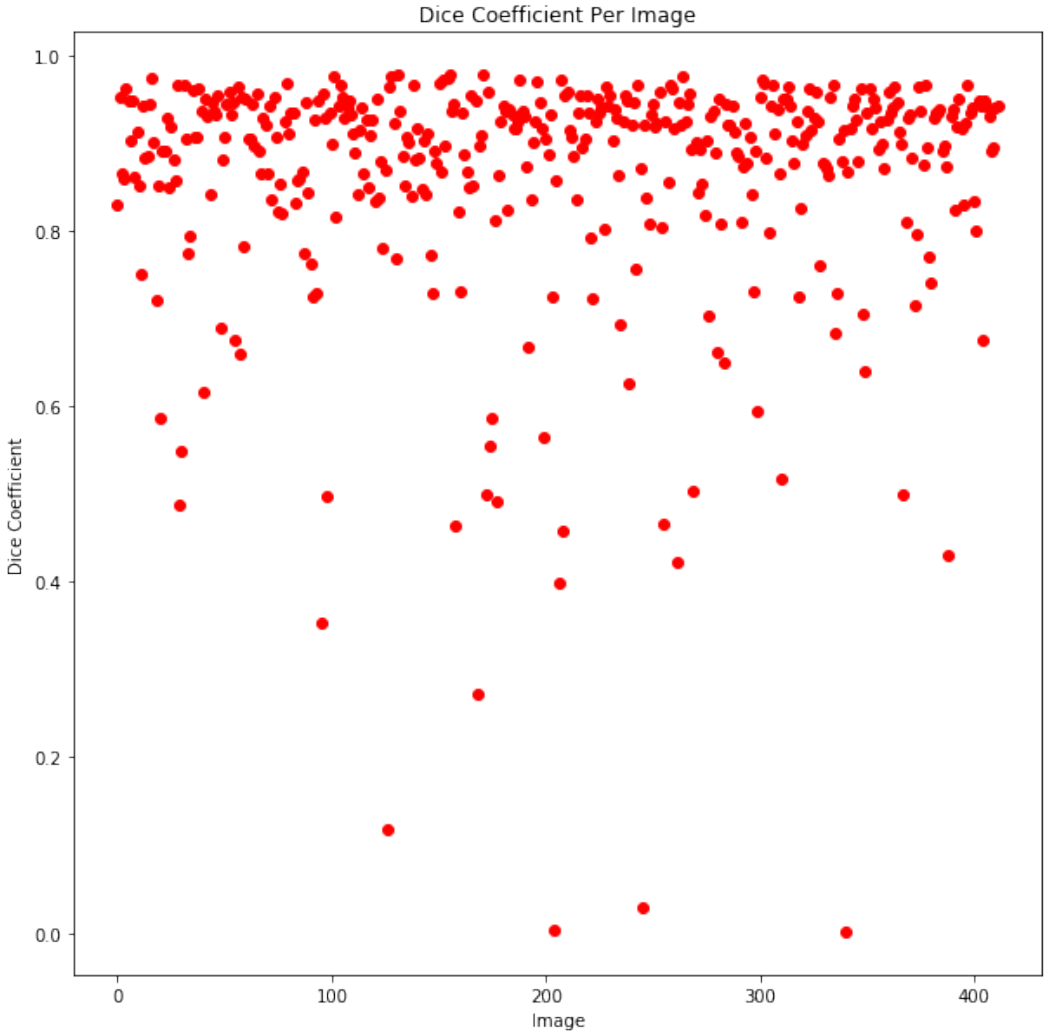

Figure 4.1: Dice coefficient for each image of the test set during Test 1. Model trained with ReLU as activation unit and Dice as loss function.

The vast majority of the worst segmentation results happened in images with lower contrast between the lesion and the background and/or in images with multiple dark areas. We took all the outliers of Test 1 and applied a alternative segmentation procedure. Using the BUSAT (Breast Ultrasound Image Analysis Toolbox)[96], the autosegment function of this toolbox computes the automatic lesion segmentation using log-Gabor filtering and texture analysis. Figure 4.2 shows an example of our segmentation and the automatic segmentation from the toolbox, and it is presented to demonstrate that those outliers were in fact complexes structures.

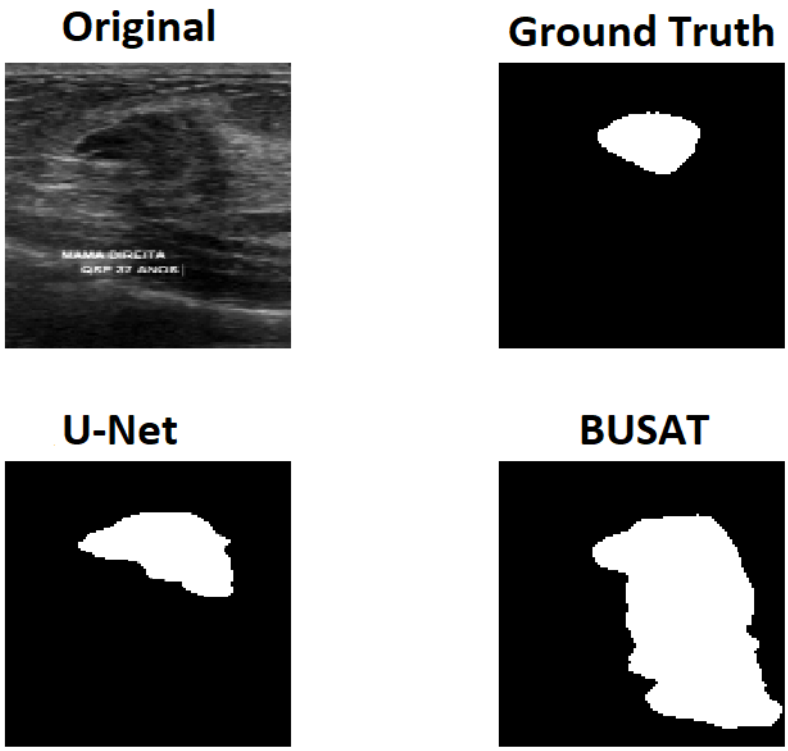

Figure 4.2: Different segmentation labels for one image of the test set. Top left: Original image; Top Right: Manually segmented; Lower Left: Test 1 output; Lower Right: BUSAT autosegment output.

4.2 SegNet

Following the previous section, the SegNet results will be presented in the same manner. The best parameters configuration is on table 4.3.

Table 4.3: Best parameters when compiling the SegNet models

Dropout Rate	**Batch Size**	**Filter Size**	**Learning Rate**
0.5	32	(5,5)	10^{-6}

Matching the U-Net parameters, the dropout rate and batch size were the same. However, the best results were scored with a larger filter size and slower learning rate. Besides, it is necessary to mention that the number of epochs for SegNet was 2000 (twice the amount used on U-Net), because in just 1000 epochs the network was not stable yet. Table 4.4 displays the results of differing functions.

Table 4.4: SegNet tests final results. The fourth column shows average dice score in the training set with standard deviation. The fifty column shows training time in hours.

#	**Activation Unit**	**Loss Function**	**Dice Score**	**Training Time (h)**
Test 5	ReLU	Dice	0.811 ± 0.154	104.45
Test 6	ReLU	Cross Entropy	0.790 ± 0.157	104.45
Test 7	PReLU	Dice	0.811 ± 0.150	127.78
Test 8	PReLU	Cross Entropy	0.790 ± 0.164	127.78

The conclusions to be made about the SegNet tests are similar to U-Net. Training time increased drastically when using PReLU instead of ReLU and no difference occurs when changing the loss function. Also, dice scores are very similar, the biggest gap being 2.1%. The Student's t-test in these configurations presents a p-value lower than 5% when comparing tests 5 and 6 (p-value = 0.002) and tests 7 and 8 (p-value = 0.0001). In this architecture, using different activation units made no impact on the dice coefficient. Once again, the dice loss function outperformed the cross entropy. Considering the slightly better or equal segmentation and the lower training time, using ReLU and dice is again the best configuration.

In Test 5, 68.2% (281) of the output labels had dice scores above average, while 7.3% (30) were outliers. The dice score for each image can be observed in figure 4.3.

Setting side by side the results of the two architectures, we can elaborate a little more. First of all, the U-Net was better than SegNet for our dataset in all aspects. U-Net was faster to train and achieved better quantitative results. Boxplots for each configuration tested in the present work are presented in figure 4.4.

While the copying of feature maps from encoder to decoder was supposed to slow the training, we were able to use a higher learning rate. When using 10^{-5} the SegNet behaved poorly. This was amended by using a slower learning rate

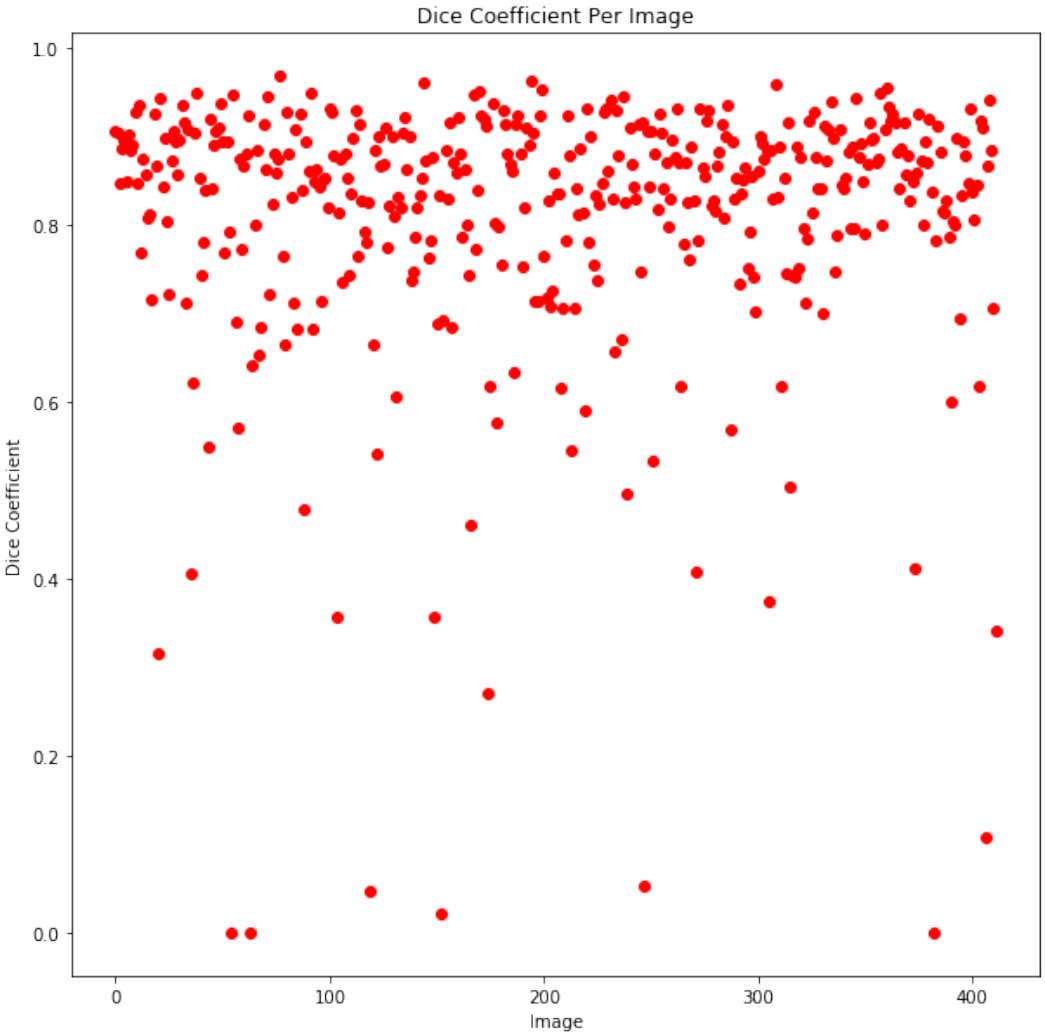

Figure 4.3: Dice coefficient for each image of the test set during Test 5. Model trained with ReLU as activation unit and Dice as loss function.

and higher number of epochs, but the training became extensively time demanding. Examples of the results obtained using U-Net's Test 1 are presented in figure 4.5.

Comparing with related works shown in Section 2.5, our work achieved dice scores on pair with the best works using U-Net and SegNet in this task. Most of the works mentioned achieved dice scores lower than ours when using the same architectures, with Zhuang et al. [16] achieving the highest SegNet score observed, 81.7% and Singh et al. [20] achieving 88.3% using U-Net.

When comparing with different techniques, the U-Net could not achieve the top results in BUS lesion segmentation. Recently, GANs [17, 18, 20] have been successfully applied to semantic segmentation, achieving dice scores higher than ours in the same task. Active contour algorithms are still able to obtain high dice scores [14, 91]. These techniques are more complex and/or slower than using U-Net, however the gain in segmentation quality is noteworthy.

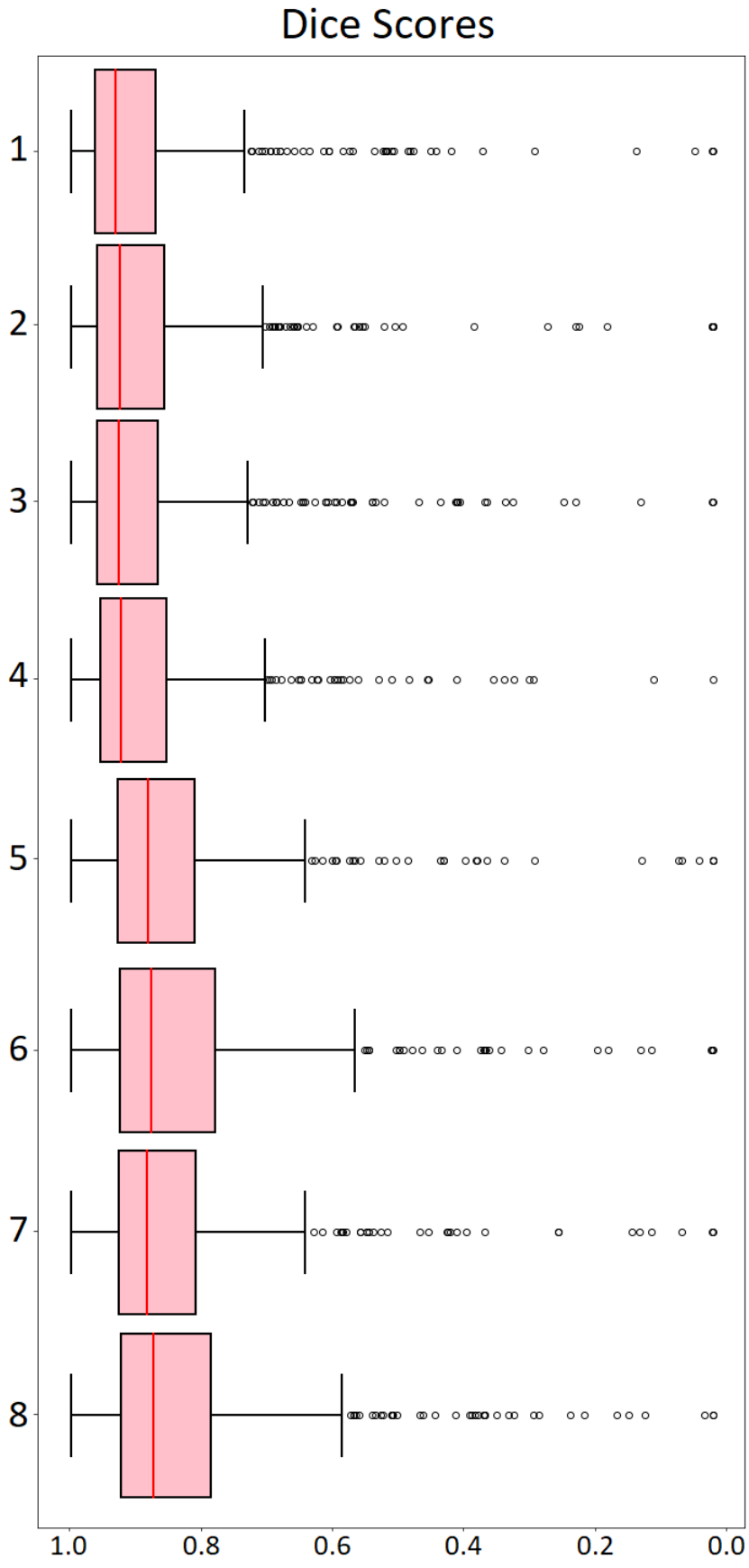

Figure 4.4: Boxplot of dice scores for each configuration in present work. Horizontal axis: dice scores. Vertical axis: test number.

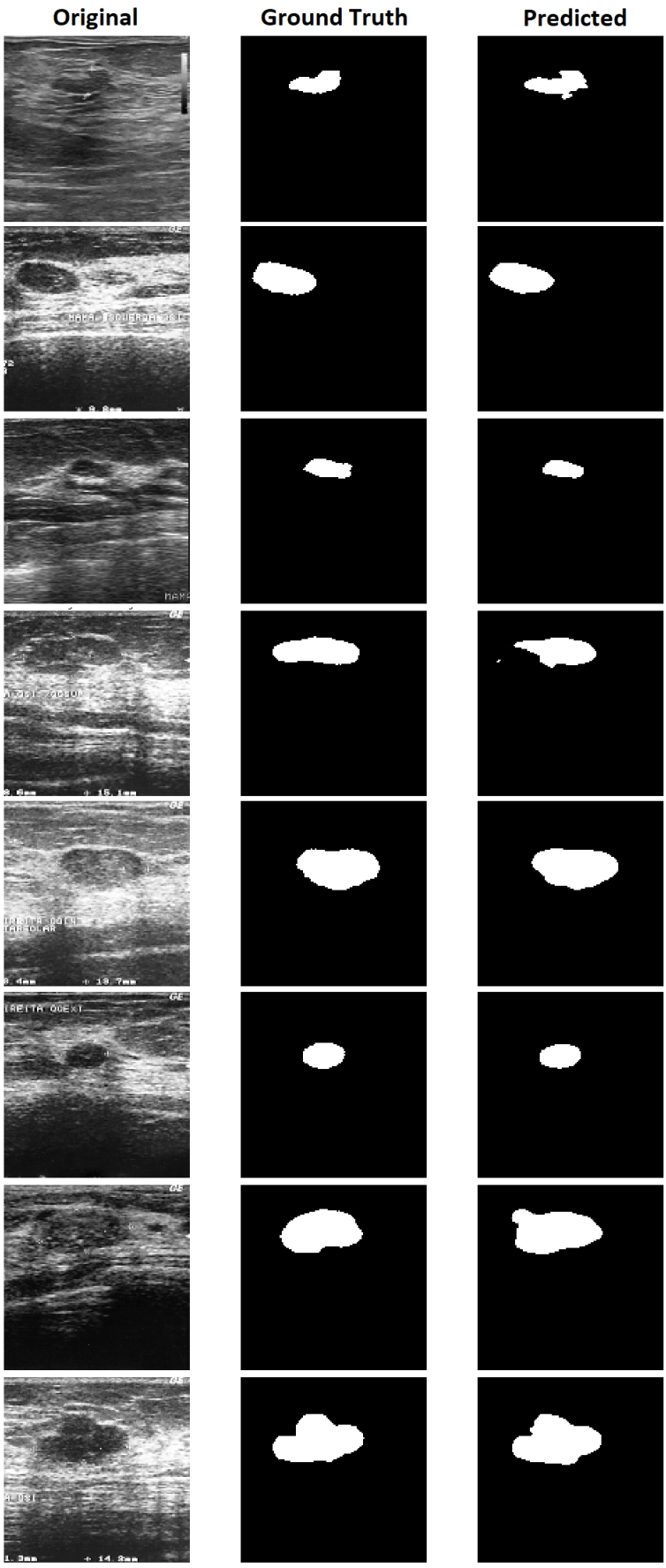

Figure 4.5: Eight examples of our results using U-Net's best configuration. In the first column are the original images, the second column shows the ground truth and the third column is the Test 1 output.

Chapter 5

Conclusion

We successfully implemented two convolutional neural networks for segmentation of breast ultrasound images lesions, with obtained accuracy on pair with most present-day works in similar tasks. Therefore, we can conclude that these two networks are able to segment ultrasound images with useful accuracy depending on their configuration. Also, comparing with different techniques, the CNNs were able to produce label outputs in a very low time.

Dice loss consistently outperformed the cross-entropy function used in the original configurations of both networks. It was expected because of the imbalance between classes in the images. Meanwhile, using parametric ReLU did not improve the results compared with the original ReLU, and the training became slower.

With the proposed architectures, we determined that the U-Net works better than the SegNet in the segmentation of our dataset. The SegNet is slower to train and achieved a lower dice score. Though U-Net is a widely used model in medical image segmentation, it has not achieved the top scores in BUS tumour segmentation. This is mainly due to the characteristics of ultrasound images. To achieve a even more accurate segmentation, the model would require more powerful feature extraction.

In spite of the progress in the last few years, this field of research is still open for future contributions. The inherent presence of noise, low contrast, sensitivity of the image-acquisition methods, and the lack of large datasets of lesions with associated labels poses as challenges in the current automatic segmentation procedure.

On a side note, we assembled two pre-processed datasets based on the original and trained the U-Net with them. The first one was denoised, while the second one was denoised and had its contrast enhanced. In both cases, the segmentation accuracy of the U-Net models did not improve in comparison with the model trained with the original dataset. This may indicate that the U-Net is learning the features on the images independently of the presence of noise or worse contrast. However, the pre-processing was done automatically and we did not had enough time to perform

substantial testing. This can be carried out in future works.

Other possibilities for future works are:

- perform a more creative data augmentation. We could use zoom, brightness changes, deformations, among other techniques. This would provide a more robust training of the network.

- using different functions. PReLU did not improve the results over ReLU, but it is possible that leaky ReLU or random ReLU present better scores. Also, using a loss function that considers more weight to pixels near the border of the lesion could help to maintain contour information.

- combine the U-Net with other techniques. Using the segmentation output of the U-Net as initialization of an active contour algorithm or using the U-Net as a generator module in a GAN are possibilities already observed in similar works.

Bibliography

[1] BRAY, F., FERLAY, J., SOERJOMATARAM, I., et al. "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries", *CA: A Cancer Journal for Clinicians*, v. 68, n. 6, pp. 394–424, nov. 2018.

[2] "Estimativa 2020: Incidência de Câncer no Brasil". Instituto Nacional do Câncer, Ministério da Saúde, 2020.

[3] FULLER, M. S., LEE, C. I., ELMORE, J. G. "Breast Cancer Screening: An Evidence-Based Update", *Medical Clinics of North America*, v. 99, n. 3, pp. 451–468, maio 2015.

[4] CALAS, M. J. G., GUTFILEN, B., PEREIRA, W. C. A. "CAD e Mamografia: por que usar essa ferramenta?" *Radiologia Brasileira*, v. 45, n. 1, pp. 46–52, jan. 2012.

[5] ONTARIO, H. Q. "Ultrasound as an Adjunct to Mammography for Breast Cancer Screening: A Health Technology Assessment", *Ontario Health Technology Assessment Series*, v. 16, n. 15, pp. 1–71, jul. 2016.

[6] DEVOLLI-DISHA, E., MANXHUKA-KËRLIU, S., YMERI, H., et al. "Comparative accuracy of mammography and ultrasound in women with breast symptoms according to age and breast density", *Bosnian Journal of Basic Medical Sciences*, v. 9, n. 2, pp. 131–136, maio 2009.

[7] DAHABREH, I. J., WIELAND, L. S., ADAM, G. P., et al. "Core Needle and Open Surgical Biopsy for Diagnosis of Breast Lesions: An Update to the 2009 Report." *Comparative Effectiveness Reviews*, v. 139, n. 14, set. 2014.

[8] ONG, M.-S., MANDL, K. D. "National Expenditure For False-Positive Mammograms And Breast Cancer Overdiagnoses Estimated At $4 Billion A Year", *Health Affairs*, v. 34, n. 4, pp. 576–583, abr. 2015.

[9] VON EULER-CHELPIN, M., LILLHOLM, M., NAPOLITANO, G., et al. "Screening mammography: benefit of double reading by breast density",

Breast cancer research and treatment, v. 171, n. 3, pp. 767–776, 2018. doi: 10.1007/s10549-018-4864-1.

[10] KIRBERGER, R. M. "Imaging artifacts in diagnostic ultrasound: a review", *Veterinary Radiology and Ultrasound*, v. 36, n. 4, pp. 297–306, jul. 1995.

[11] XIAN, M., ZHANG, Y., CHENG, H. D., et al. "Automatic breast ultrasound image segmentation: A survey", *Pattern Recognition*, v. 79, n. 1, pp. 340–355, jul. 2018.

[12] DA SILVA, F. H. S. *Deep Learning for Corpus Callosum Segmentation in Brain Magnetic Reessonance Images*. M.Sc. dissertation, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brasil, 2018.

[13] XIE, X., SHI, F., NIU, J., et al. "Breast Ultrasound Image Classification and Segmentation Using Convolutional Neural Networks", *Lecture Notes in Computer Science*, v. Advances in Multimedia Information Processing – PCM 2018, pp. 200–211, 2018.

[14] HU, Y., GUO, Y., WANG, Y., et al. "Automatic tumor segmentation in breast ultrasound images using a dilated fully convolutional network combined with an active contour model", *Medical Physics*, out. 2018.

[15] ALMAJALID, R., SHAN, J., DU, Y., et al. "Development of a Deep-Learning-Based Method for Breast Ultrasound Image Segmentation". In: *17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018, Orlando, FL, USA, December 17-20*, pp. 1103–1108, 2018. doi: 10.1109/ICMLA.2018.00179.

[16] ZHUANG, Z., LI, N., RAJ, A. N. J., et al. "An RDAU-NET model for lesion segmentation in breast ultrasound images", *PLOS ONE*, v. 14, n. 8, pp. 1–23, 08 2019. doi: 10.1371/journal.pone.0221535. Disponível em: <`https://doi.org/10.1371/journal.pone.0221535`>.

[17] NEGI, A., RAJ, A. N. J., NERSISSON, R., et al. "RDA-UNET-WGAN: An Accurate Breast Ultrasound Lesion Segmentation Using Wasserstein Generative Adversarial Networks", *Arabian Journal for Science and Engineering*, 2020. doi: 10.1007/s13369-020-04480-z.

[18] HAN, L., HUANG, Y., DOU, H., et al. "Semi-supervised segmentation of lesion from breast ultrasound images with attentional generative adversarial network", *Computer Methods and Programs in Biomedicine*, v. 189, pp. 105275, 2020. doi: 10.1016/j.cmpb.2019.105275.

[19] GHOSH, D., KUMAR, A., GHOSAL, P., et al. "Breast Lesion Segmentation in Ultrasound Images Using Deep Convolutional Neural Networks". In: *2020 IEEE Calcutta Conference (CALCON)*, pp. 318–322, 2020.

[20] SINGH, V. K., RASHWAN, H. A., ABDEL-NASSER, M., et al. "An Efficient Solution for Breast Tumor Segmentation and Classification in Ultrasound Images Using Deep Adversarial Learning". 2019.

[21] HORNBERG, J. J., BRUGGEMAN, F. J., WESTERHOFF, H. V., et al. "Cancer: A Systems Biology disease", *Biosystems*, v. 83, n. 2–3, pp. 81–90, fev. 2006.

[22] HARRINGTON, K. J. "Biology of cancer", *Medicine*, v. 12, pp. 689–692, dez. 2011.

[23] BARRERA, M. "How cancer starts, grows and spreads", *Canadian Cancer Society <http://www.cancer.ca/en/cancer-information/cancer-101/...what-is-cancer/how-cancer-starts-grows-and-spreads>*, nov. 2018.

[24] *Informed Health Online [Internet]. Cologne, Germany: Institute for Quality and Efficiency in Health Care (IQWiG); 2006-. How do cancer cells grow and spread? 2013 Nov 6 [Updated 2016 Sep 21]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK279410/*, nov. 2018.

[25] *American Cancer Society. USA; 2018. What Is Breast Cancer? 2017 Aug 1 [Updated 2017 Sep 21]. Available from: https://www.cancer.org/cancer/breast-cancer/about/what-is-breast-cancer.html*, nov. 2018.

[26] SHARMA, G. N., DAVE, R., SANADYA, J., et al. "Various types and management of breast cancer: an overview", *Journal of Advanced Pharmaceutical Technology and Research*, v. 1, n. 2, pp. 109–126, abr. 2010.

[27] JEMAL, A., CLEGG, L. X., WARD, E., et al. "Annual report to the nation on the status of cancer, 1975-2001, with a special feature regarding survival", *Cancer*, v. 101, n. 1, pp. 3–27, jul. 2004.

[28] MA, I., FISHELL, F., WRIGHT, B., et al. "Case-control study of factors associated with failure to detect breast cancer by mammography", *Journal of National Cancer Institute*, v. 84, n. 10, pp. 781–785, maio 1992.

[29] PELLERITO, J., POLAK, J. F. *Introduction to Vascular Ultrasonography: Expert Consult - Online and Print*. 6 ed. New York, Elsevier Health Sciences, 2012.

[30] FISH, P. *Physics and Instrumentation of Diagnostic Medical Ultrasound*. Illustrated ed. Michigan, Wiley, 1990.

[31] D'ASTOUS, F. T., FOSTER, F. S. "Frequency dependence of ultrasound attenuation and backscatter in breast tissue." *Ultrasound in Medicine and Biology*, v. 12, n. 10, pp. 795–808, maio 1986.

[32] CARLSEN, E. N. "Ultrasound physics for the physician a brief review." *Journal of Clinical Ultrasound*, v. 3, n. 1, pp. 69–75, mar. 1975.

[33] GOLDBERG, B. B. *Medical Diagnostic Ultrasound: A Retrospective on Its 40th Anniversary*. Rochester, NY, Kodak Health Sciences, 1988.

[34] STUDYGROUP, W. H. O. *Training in diagnostic ultrasound: essentials, principles and standards.* World Health Organization 875, 1998.

[35] TEH, W., WILSON, A. R. M. "The role of ultrasound in breast cancer screening. A consensus statement by the European Group for breast cancer screening", *European Journal of Cancer*, v. 34, n. 4, pp. 449–450, mar. 1998.

[36] MANDELSON, M., OESTREICHER, N., PORTER, P. L., et al. "Breast density as a predictor of mammographic detection: comparison of interval- and screen-detected cancers", *Journal of the National Cancer Institute*, v. 92, n. 13, pp. 1081–1087, jul. 2000.

[37] GEISEL, J., RAGHU, M., HOOLEY, R. "The Role of Ultrasound in Breast Cancer Screening: The Case for and Against Ultrasound", *Seminars in Ultrasound, CT and MRI*, v. 39, n. 1, pp. 25–34, fev. 2018.

[38] HELMUT, M., MENDELSON, E. B. *The Practice of Breast Ultrasound: Techniques, Findings, Differential Diagnosis*. Chicago, Il, Thieme, 2008.

[39] SEHGAL, C. M., WEINSTEIN, S. P., ARGER, P. H., et al. "A review of breast ultrasound", *Journal of Mammary Gland Biology and Neoplasia*, v. 11, n. 2, pp. 113–123, abr. 2006.

[40] ALVARENGA, A. V. *Quantificação das características morfométricas e de textura de tumores de mama em imagens por ultrassom e a avaliação da sua contribuição diagnóstica*. Tese de D.Sc., COPPE/UFRJ, Rio de Janeiro, RJ, Brasil, 2005.

[41] TAYLOR-PHILLIPS, S., JENKINSON, D., STINTON, C., et al. "Double Reading in Breast Cancer Screening: Cohort Evaluation in the CO-OPS Trial", *Radiology*, v. 287, n. 3, pp. 749–757, jun. 2018.

[42] CHEN, C.-M., CHOU, Y.-H., HAN, K.-C., et al. "Breast Lesions on Sonograms: Computer-aided Diagnosis with Nearly Setting-Independent Features and Artificial Neural Networks", *Radiology*, v. 226, n. 2, pp. 504–514, fev. 2003.

[43] CHENG, H. D., SHAN, J., JU, W., et al. "Automated breast cancer detection and classification using ultrasound images: A survey", *Pattern Recognition*, v. 43, n. 1, pp. 299–317, jan. 2010.

[44] CHENG, H., SHI, X. "A simple and effective histogram equalization approach to image enhancement", *Digital Signal Processing*, v. 14, n. 2, pp. 158–170, mar. 2004.

[45] ZUIDERVELD, K. "Contrast limited adaptive histogram equalization". In: Heckbert, P. S. (Ed.), *Graphics Gems IV*, illustrated ed., San Diego, CA, Academic Press Professional, 1994.

[46] AWAD, J., ABDEL-GALIL, T. K., SALAMA, M. M. A., et al. "Prostate's boundary detection in transrectal ultrasound images using scanning technique", *CCECE 2003 - Canadian Conference on Electrical and Computer Engineering*, maio 2003.

[47] GUO, Y., CHENG, H. D., HUANG, J., et al. "Breast ultrasound image enhancement using fuzzy logic", *Ultrasound in Medicine Biology*, v. 32, n. 2, pp. 237–247, fev. 2006.

[48] WANG, Y., WANG, H., GUO, Y., et al. "Novel Computer-Aided Diagnosis Algorithms on Ultrasound Image: Effects on Solid Breast Masses Discrimination", *Journal of Digital Imaging*, v. 23, n. 5, pp. 581—591, out. 2010.

[49] FLORES, W. G., PEREIRA, W. C. A. "A contrast enhancement method for improving the segmentation of breast lesions on ultrasonography", *Computers in Biology and Medicine*, v. 80, n. 1, pp. 14–23, jan. 2017.

[50] HORSCH, K., GIGER, M. L., VENTA, L. A., et al. "Automatic segmentation of breast lesions on ultrasound", *Medical Physics*, v. 28, n. 8, pp. 1652–1659, ago. 2001.

[51] RUI, W., JIANGLI, L., DEYU, L., et al. "Edge Enhancement and Filtering of Medical Ultrasonic Images Using a Hybrid Method", *International Conference on Bioinformatics and Biomedical Engineering*, , n. 1, jul. 2007.

[52] GUPTA, S., KAUR, L., CHAUHAN, R., et al. "A versatile technique for visual enhancement of medical ultrasound images", *Digital Signal Processing*, v. 17, n. 3, pp. 542–560, maio 2007.

[53] ADAM, D., BEILIN-NISSAN, S., Z. FRIEDMAN, V. B. "The combined effect of spatial compounding and nonlinear filtering on the speckle reduction in ultrasound images", *Ultrasonics*, v. 44, n. 2, pp. 166–181, nov. 2005.

[54] PERONA, P., MALIK, J. "Scale-Spacing and Edge Detection Using Anisotropic Diffusion", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 12, n. 7, pp. 629–639, jul. 1990.

[55] RODTOOK, A., S.MAKHANOV, S. "Multi-feature gradient vector flow snakes for adaptive segmentation of the ultrasound images of breast cancer", *Journal of Visual Communication and Image Representation*, v. 24, n. 8, pp. 1414–1430, nov. 2013.

[56] CHEN, D.-R., HSIAO, Y.-H. "Computer-aided Diagnosis in Breast Ultrasound", *Journal of Medical Ultrasound*, v. 16, n. 1, pp. 46–56, 2008.

[57] CHANG, R. F., WU, W.-J., TSENG, C.-C., et al. "3-D snake for US in margin evaluation for malignant breast tumor excision using mammotome", *IEEE Transactions on Information Technology in Biomedicine*, v. 7, n. 3, pp. 197–201, set. 2003.

[58] H. D. CHENG, L. HU, J. T. L. S. "A Novel Markov Random Field Segmentation Algorithm and Its Application to Breast Ultrasound Image Analysis", *International Conference on Computer Vision, Pattern Recognition and Image Processing*, , n. 6, 2005.

[59] CHEN, D.-R., CHANG, R.-F., KUO, W.-J., et al. "Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks", *Ultrasound in Medicine and Biology*, v. 28, n. 10, pp. 1301–1310, out. 2002.

[60] LÓPEZ, J. H. *Sintonización automática de la red neuronal de pulso acoplado mediante evolución diferencial para la segmentación de imágenes.* M.Sc thesis, Instituto Politécnico Nacional, Ciudad Victoria, Tamaulipas, México, 2016.

[61] FLORES, W. G., PEREIRA, W. C. A., INFANTOSI, A. F. C. "Improving Classification Performance of Breast Lesions on Ultrasonography", *Pattern Recognition*, v. 48, n. 4, pp. 1121–1132, abr. 2015.

[62] PEREIRA, W. C. A., ALVARENGA, A. V., INFANTOSI, A. F. C., et al. "A non-linear morphometric feature selection approach for breast tumor contour from ultrasonic images", *Computers in Biology and Medicine*, v. 40, n. 11, pp. 912–918, nov. 2010.

[63] ALVARENGA, A. V., PEREIRA, W. C. A., INFANTOSI, A. F. C., et al. "Assessing the combined performance of texture and morphological parameters in distinguishing breast tumors in ultrasound images", *Medical Physics*, v. 39, n. 12, pp. 7350–7358, dez. 2012.

[64] NAYEEM, M. A. R., JOADDER, M. A. M., SHETU, S. A., et al. "Feature selection for breast cancer detection from ultrasound images". In: *2014 International Conference on Informatics, Electronics Vision (ICIEV)*, pp. 1–6, 2014.

[65] CHOU, Y. H., TIU, C. M., HUNG, G. S., et al. "Stepwise Logistic Regression Analysis of Tumor Contour Features for Breast Ultrasound Diagnosis", *Ultrasound in medicine and biology*, v. 27, n. 11, pp. 1493–1498, 2001. doi: 10.1016/s0301-5629(01)00466-5.

[66] ALVARENGA, A. V., INFANTOSI, A. F. C., PEREIRA, W. C. A., et al. "Assessing the performance of morphological parameters in distinguishing breast tumors on ultrasound images", *Medical Engineering and Physics*, v. 32, n. 1, pp. 49–56, jan. 2010.

[67] JOO, S., YANG, Y. S., MOON, W. K., et al. "Computer-aided diagnosis of solid breast nodules: use of an artificial neural network based on multiple sonographic features", *IEEE transactions on medical imaging*, v. 23, n. 10, pp. 1292–1300, out. 2004.

[68] KUO, W.-J., CHANG, R.-F., CHEN, D.-R., et al. "Data mining with decision trees for diagnosis of breast tumor in medical ultrasonic images", *Breast Cancer Research and Treatment*, v. 66, n. 1, pp. 51–57, mar. 2001.

[69] HUANG, Y.-L., CHEN, D.-R. "Support vector machines in sonography: Application to decision making in the diagnosis of breast cancer", *Journal of Clinical Imaging*, v. 29, n. 3, pp. 179–184, maio 2005.

[70] KUO, W.-J., CHANG, R.-F., LEE, C. C., et al. "Retrieval technique for the diagnosis of solid breast tumors on sonogram", *Ultrasound in Medicine and Biology*, v. 28, n. 7, pp. 903—909, jul. 2002.

[71] HORSCH, K., GIGER, M. L., VYBORNY, C. J., et al. "Performance of computer-aided diagnosis in the interpretation of lesions on breast sonography", *Academic Radiology*, v. 11, n. 3, pp. 272—280, mar. 2004.

[72] HORSCH, K., GIGER, M. L., VYBORNY, C. J., et al. "Classification of Breast Lesions with Multimodality Computer-aided Diagnosis: Observer Study Results on an Independent Clinical Data Set", *Radiology*, v. 240, n. 2, pp. 357—368, ago. 2006.

[73] GRUSZAUSKAS, N. P., DRUKKER, K., GIGER, M. L., et al. "Performance of Breast Ultrasound Computer-aided Diagnosis. Dependence on Image Selection", *Academic Radiology*, v. 15, n. 10, pp. 1234—1245, nov. 2008.

[74] CHOI, J.-H., KANG, B. J., JI EUN BAEK AND, H. S. L., et al. "Application of computer-aided diagnosis in breast ultrasound interpretation: improvements in diagnostic performance according to reader experience", *Ultrasonography*, v. 37, n. 3, pp. 217—225, jul. 2018.

[75] SHAN, J. *A Fully Automatic Segmentation Method for Breast Ultrasound Images*. Tese de D.Sc., Utah State University, Logan, UT, 2011.

[76] BROWN, A. *Introduction to Object Detection and Image Segmentation*. Presentation, NVidia, 2011.

[77] BROSNAN, T., SUN, D.-W. "Improving quality inspection of food products by computer vision—a review", *Journal of Food Engineering*, v. 61, n. 1, pp. 3—16, jan. 2004.

[78] RONNEBERGER, O., FISCHER, P., BROX, T. "U-Net: Convolutional Networks for Biomedical Image Segmentation", *arXiv*, maio 2015.

[79] LECUN, Y., BENGIO, Y., HINTON, G. "Deep Learning", *Nature*, v. 521, n. 7553, pp. 436–444, maio 2015.

[80] ZHANG, X., CHEN, M. H., QIN, Y. "NLP-QA Framework Based on LSTM-RNN". In: *2018 2nd International Conference on Data Science and Business Analytics (ICDSBA)*, pp. 307–311, Sep. 2018. doi: 10.1109/ICDSBA.2018.00065.

[81] KRIZHEVSKY, A., SUTSKEVER, I., HINTON, G. E. "ImageNet Classification with Deep Convolutional Neural Networks", *Advances in neural information processing systems*, v. 25, n. 2, jan. 2012.

[82] SIMONYAN, K., ZISSERMAN, A. "Very deep convolutional networks for large-scale image recognition", *arXiv*, set. 2014.

[83] SZEGEDY, C., LIU, W., JIA, Y., et al. "Going deeper with convolutions", *arXiv*, set. 2014.

[84] HE, K., ZHANG, X., REN, S., et al. "Deep residual learning for image recognition", *arXiv*, dez. 2015.

[85] BADRINARAYANAN, V., KENDALL, A., CIPOLLA, R. "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation", *arXiv*, out. 2016.

[86] DESHPANDE, A. "A Beginner's Guide To Understanding Convolutional Neural Networks", *July 2016. Acessed on November 2018. Available from: https://adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/.*

[87] KARPATHY, A. "CS231n Convolutional Neural Networks for Visual Recognition", *Stanford CS Class. Acessed on November 2018. Available from: http://cs231n.github.io/convolutional-networks/.*

[88] SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., et al. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting", *Journal of Machine Learning Research*, v. 15, pp. 1929–1958, jun. 2014.

[89] IOFFE, S., SZEGEDY, C. "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift", *arXiv*, mar. 2015.

[90] XIAN, M., ZHANG, Y., CHENG, H. D., et al. "A Benchmark for Breast Ultrasound Image Segmentation (BUSIS)". 2018.

[91] KRITI, VIRMANI, J., AGARWAL, R. "Assessment of despeckle filtering algorithms for segmentation of breast tumours from ultrasound images", *Biocybernetics and Biomedical Engineering*, v. 39, n. 1, pp. 100 – 121, 2019. ISSN: 0208-5216. doi: https://doi.org/10.1016/j.bbe.2018.10.002. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0208521618300652>.

[92] VAKANSKI, A., XIAN, M., FREER, P. "Attention Enriched Deep Learning Model for Breast Tumor Segmentation in Ultrasound Images". 2019.

[93] LU, L., SHIN, Y., SU, Y., et al. “Dying ReLU and Initialization: Theory and Numerical Examples”. 2019.

[94] XU, B., WANG, N., CHEN, T., et al. “Empirical Evaluation of Rectified Activations in Convolutional Network”. 2015.

[95] SUDRE, C. H., LI, W., VERCAUTEREN, T., et al. “Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations”, *Lecture Notes in Computer Science*, p. 240–248, 2017. ISSN: 1611-3349. doi: 10.1007/978-3-319-67558-9_28. Disponível em: <http://dx.doi.org/10.1007/978-3-319-67558-9_28>.

[96] RODRÍGUEZ-CRISTERNA, A., GÓMEZ-FLORES, W., DE ALBUQUERQUE-PEREIRA, W. C. “BUSAT: A MATLAB Toolbox for Breast Ultrasound Image Analysis”. In: Carrasco-Ochoa, J. A., Martínez-Trinidad, J. F., Olvera-López, J. A. (Eds.), *Pattern Recognition*, pp. 268–277, Cham, 2017. Springer International Publishing. ISBN: 978-3-319-59226-8.