



## **ANALYZING THE BREAST TISSUE IN MAMMOGRAMS USING DEEP LEARNING**

**Nasibeh Saffari Tabalvandani**

**ADVERTIMENT.** L'accés als continguts d'aquesta tesi doctoral i la seva utilització ha de respectar els drets de la persona autora. Pot ser utilitzada per a consulta o estudi personal, així com en activitats o materials d'investigació i docència en els termes establerts a l'art. 32 del Text Refós de la Llei de Propietat Intel·lectual (RDL 1/1996). Per altres utilitzacions es requereix l'autorització prèvia i expressa de la persona autora. En qualsevol cas, en la utilització dels seus continguts caldrà indicar de forma clara el nom i cognoms de la persona autora i el títol de la tesi doctoral. No s'autoritza la seva reproducció o altres formes d'explotació efectuades amb finalitats de lucre ni la seva comunicació pública des d'un lloc aliè al servei TDX. Tampoc s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX (framing). Aquesta reserva de drets afecta tant als continguts de la tesi com als seus resums i índexs.

**ADVERTENCIA.** El acceso a los contenidos de esta tesis doctoral y su utilización debe respetar los derechos de la persona autora. Puede ser utilizada para consulta o estudio personal, así como en actividades o materiales de investigación y docencia en los términos establecidos en el art. 32 del Texto Refundido de la Ley de Propiedad Intelectual (RDL 1/1996). Para otros usos se requiere la autorización previa y expresa de la persona autora. En cualquier caso, en la utilización de sus contenidos se deberá indicar de forma clara el nombre y apellidos de la persona autora y el título de la tesis doctoral. No se autoriza su reproducción u otras formas de explotación efectuadas con fines lucrativos ni su comunicación pública desde un sitio ajeno al servicio TDR. Tampoco se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR (framing). Esta reserva de derechos afecta tanto al contenido de la tesis como a sus resúmenes e índices.

**WARNING.** Access to the contents of this doctoral thesis and its use must respect the rights of the author. It can be used for reference or private study, as well as research and learning activities or materials in the terms established by the 32nd article of the Spanish Consolidated Copyright Act (RDL 1/1996). Express and previous authorization of the author is required for any other uses. In any case, when using its content, full name of the author and title of the thesis must be clearly indicated. Reproduction or other forms of for profit use or public communication from outside TDX service is not allowed. Presentation of its content in a window or frame external to TDX (framing) is not authorized either. These rights affect both the content of the thesis and its abstracts and indexes.



# Analyzing the breast tissue in mammograms using deep learning



**Nasibeh Saffari**

*DOCTORAL THESIS*  
2022

# Analyzing the breast tissue in mammograms using deep learning

DOCTORAL THESIS

Author:

Nasibeh Saffari

Advisors:

Dr. Blas Herrera Gomez

Prof. Dr. Domenec Savi Puig Valls

*Departament d'Enginyeria Informàtica i Matemàtiques (DEIM)*



UNIVERSITAT ROVIRA I VIRGILI

Tarragona

2022



UNIVERSITAT ROVIRA I VIRGILI

*Departament d'Enginyeria Informatica i Matematiques (DEIM)*  
*Av. Paisos Catalans, 26*  
*43007 Tarragona, Spain*

We STATE that the present study, entitled "Breast density analysis in the mammograms," presented by Nasibeh Saffari, for the award of the degree of Doctor, has been carried out under our supervision at the Department d'Enginyeria *Informatica i Matematiques* (DEIM).

Tarragona, 23<sup>th</sup> January 2022.

Doctoral Thesis Supervisors,

Prof. Dr. Domenec Savi Puig Valls

Dr. Blas Herrera Gomez

تقدیم بہ پدر بزرگوار و مادر عزیزم

**Dedicated to My dear father, mother, and my brother**

## Abstract

Mammographic breast density (MBD) reflects the amount of fibroglandular breast tissue area that appears white and bright on mammograms, commonly referred to as breast percent density (PD%). MBD is a risk factor for breast cancer and a risk factor for masking tumors. However, accurate MBD estimation with visual assessment is still a challenge due to faint contrast and significant variations in background fatty tissues in mammograms. In addition, correctly interpreting mammogram images requires highly trained medical experts: it is difficult, time-consuming, expensive, and error-prone. Nevertheless, dense breast tissue can make it harder to identify breast cancer and be associated with an increased risk of breast cancer. For example, it has been reported that women with a high breast density compared to women with a low breast density have a four- to six-fold increased risk of developing the disease.

The primary key of breast density computing and breast density classification is to detect the dense tissues in the mammographic images correctly. Many methods have been proposed for breast density estimation; however, most are not automated. Besides, they have been badly affected by low signal-to-noise ratio and variability of density in appearance and texture. It would be more helpful to have a computer-aided diagnosis (CAD) system to assist the doctor analyze and diagnosing it automatically. Current development in deep learning methods motivates us to improve current breast density analysis systems.

The main focus of the present thesis is to develop a system for automating the breast density analysis (such as; breast density segmentation(BDS), breast density percentage (BDP), and breast density classification ( BDC)), using deep learning

techniques and applying it on the temporal mammograms after treatment for analyzing the breast density changes to find a risky and suspicious patient.

Firstly, we applied the Vgg-Segnet, FCN-8, FCN-32, and cGAN model with MSE loss function for breast density segmentation on the Inbreast dataset. The cGAN model with MSE loss function yields the best result to segment the dense regions in mammographic images with overall recall, precision, and F-score around 95%, 92%, and 93%, respectively, outperforming the common semantic segmentation methods (FCN and SegNet).

Second, the conditional Generative Adversarial Networks (cGAN) with dice loss function was applied to segment dense tissues in mammograms; two methods, including thresholding and Convolutional neural network (CNN), were applied used to perform the breast density classification.

Finally, we proposed the new method for computing the breast density changes after treatment by analyzing the combination of the whole breast density and mass region density which is calculated by applying the conditional Generative Adversarial Networks (cGAN) with dice loss function to segment the breast density and mass region density after registration in the three years temporal mammograms.

Overall, this thesis presents the automated deep learning method for breast density analysis; This thesis 's findings are promising and show that the proposed techniques can produce a clinically helpful computer-aided tool for breast density analysis by digital mammography.

**Keywords:** Breast cancer, Breast density estimation, Deep learning, Generative Adversarial Networks, Mammograms.

## Resumen (Spanish)

La densidad mamográfica de la mama (MBD) refleja la cantidad de área fibroglandular del tejido mamario que aparece blanca y brillante en las mamografías, comúnmente conocida como densidad porcentual de la mama (PD%). El MBD es un factor de riesgo para el cáncer de mama y un factor de riesgo para enmascarar tumores. Sin embargo, la estimación precisa de la DMO con evaluación visual sigue siendo un reto debido al contraste débil y a las variaciones significativas en los tejidos grasos de fondo en las mamografías. Además, la interpretación correcta de las imágenes de mamografía requiere de expertos médicos altamente capacitados: Es difícil, laborioso, caro y propenso a errores. Sin embargo, el tejido mamario denso puede dificultar la identificación del cáncer de mama y asociarse con un mayor riesgo de cáncer de mama. Por ejemplo, se ha informado que las mujeres con una alta densidad mamaria en comparación con las mujeres con una densidad mamaria baja tienen un riesgo de cuatro a seis veces mayor de desarrollar la enfermedad.

La clave principal de la computación de densidad de mama y la clasificación de densidad de mama es detectar correctamente los tejidos densos en las imágenes mamográficas. Se han propuesto muchos métodos para la estimación de la densidad mamaria; sin embargo, la mayoría de ellos no están automatizados. Además, se han visto gravemente afectados por la baja relación señal-ruido y la variabilidad de la densidad en apariencia y textura.

Sería más útil disponer de un sistema de diagnóstico asistido por ordenador (CAD) para ayudar al médico a analizarlo y diagnosticarlo automáticamente. El desarrollo actual



de métodos de aprendizaje profundo nos motiva a mejorar los sistemas actuales de análisis de densidad mamaria.

El enfoque principal de la presente tesis es desarrollar un sistema para automatizar el análisis de densidad de la mama ( tal como; Segmentación de densidad de mama (BDS), porcentaje de densidad de mama (BDP) y clasificación de densidad de mama (BDC)), utilizando técnicas de aprendizaje profundo y aplicándola en las mamografías temporales después del tratamiento para analizar los cambios de densidad de mama para encontrar un paciente peligroso y sospechoso.

En primer lugar, se aplicaron los modelos VGG-Segnet, FCN-8, FCN-32 y cGAN con función de pérdida de MSE para la segmentación de la densidad de mama en el conjunto de datos Inmamet. El modelo cGAN con función de pérdida de MSE proporciona el mejor resultado para segmentar las regiones densas en imágenes mamográficas con recuperación general, precisión y puntuación F en torno al 95%, 92% y 93%, respectivamente, superando los métodos de segmentación semántica comunes (FCN y SegNet).

En segundo lugar, se aplicó la función condicional de Redes Advertisarias Generativas (cGAN) con función de pérdida de datos para segmentar tejidos densos en mamografías; se aplicaron dos métodos, incluyendo la red neural de umbral y convolucional (CNN), para realizar la clasificación de la densidad de la mama.

Por último, se propuso el nuevo método para calcular los cambios en la densidad de la mama después del tratamiento mediante el análisis de la combinación de la densidad de la mama completa y la densidad de la región de masa, que se calcula aplicando las Redes de Advertimiento Generativo Condicional (cGAN) con función de pérdida de datos

para segmentar la densidad de la mama y la densidad de la región de masa después del registro en los mamogramas temporales de tres años.

En general, esta tesis presenta el método automatizado de aprendizaje profundo para el análisis de la densidad mamaria; los resultados de este estudio son prometedores y muestran que las técnicas propuestas pueden producir una herramienta clínicamente útil asistida por ordenador para el análisis de la densidad mamaria mediante mamografía digital.

## Resum (Catalan)

La densitat mamogràfica de la mama (MBD) reflecteix la quantitat d'àrea fibroglandular del teixit mamari que apareix blanca i brillant a les mamografies, comunament coneguda com a densitat percentual de la mama (PD%). El MBD és un factor de risc per al càncer de mama i un factor de risc per emmascarar tumors. Tot i això, l'estimació precisa de la DMO amb avaluació visual continua sent un repte a causa del contrast feble i de les variacions significatives en els teixits grassos de fons en les mamografies. A més, la interpretació correcta de les imatges de mamografia requereix experts mèdics altament capacitats: És difícil, laboriós, car i propens a errors. No obstant això, el teixit mamari dens pot dificultar la identificació del càncer de mama i associar-se amb un risc més gran de càncer de mama. Per exemple, s'ha informat que les dones amb una alta densitat mamària en comparació amb les dones amb una densitat mamària baixa tenen un risc de quatre a sis vegades més gran de desenvolupar la malaltia.

La clau principal de la computació de densitat de mama i la classificació de densitat de mama és detectar correctament els teixits densos a les imatges mamogràfiques. S'han proposat molts mètodes per estimar la densitat mamària; no obstant això, la majoria no estan automatitzats. A més, s'han vist greument afectats per la baixa relació senyal-soroll i la variabilitat de la densitat en aparença i textura.

Seria més útil tenir un sistema de diagnòstic assistit per ordinador (CAD) per ajudar el metge a analitzar-lo i diagnosticar-lo automàticament. El desenvolupament actual de mètodes d'aprenentatge profund ens motiva a millorar els sistemes actuals d'anàlisi de densitat mamària.

L'enfocament principal de la present tesi és desenvolupar un sistema per automatitzar l'anàlisi de densitat de la mama ( tal com; Segmentació de densitat de mama (BDS), percentatge de densitat de mama (BDP) i classificació de densitat de mama (BDC) ), utilitzant tècniques d'aprenentatge profund i aplicant-la a les mamografies temporals després del tractament per analitzar els canvis de densitat de mama per trobar un pacient perillós i sospitós.

En primer lloc, es van aplicar els models VGG-Segnet, FCN-8, FCN-32 i cGAN amb funció de pèrdua de MSE per a la segmentació de la densitat de mama al conjunt de dades Inmamet. El model cGAN amb funció de pèrdua de MSE proporciona el millor resultat per segmentar les regions denses en imatges mamogràfiques amb recuperació general, precisió i puntuació F al voltant del 95%, 92% i 93%, respectivament, superant els mètodes de segmentació semàntica comuns (FCN i SegNet).

En segon lloc, es va aplicar la funció condicional de Xarxes Advertisàries Generatives (cGAN) amb funció de pèrdua de daus per segmentar teixits denses en mamografies; es van aplicar dos mètodes, incloent-hi la xarxa neural de llindar i convolucional (CNN), per realitzar la classificació de la densitat de la mama.

Finalment, es va proposar el nou mètode per calcular els canvis en la densitat de la mama després del tractament mitjançant l'anàlisi de la combinació de la densitat de la mama completa i la densitat de la regió de massa, que es calcula aplicant les xarxes de Advertiment Generatiu Condicional (cGAN) amb funció de pèrdua de daus per segmentar la densitat de la mama i la densitat de la regió de massa després del registre en els mamogrames temporals de tres anys.

En general aquesta tesi presenta el mètode automatitzat d'aprenentatge profund per a l'anàlisi de la densitat mamària; els resultats d'aquest estudi són prometedors i mostren

que les tècniques proposades poden produir una eina clínicament útil assistida per ordinador per a l'anàlisi de la densitat mamària mitjançant mamografia digital.

## Acknowledgments

I would like to express my gratitude to my supervisors Prof. Dr. Domenech Savi Puig Valls and Dr. Blas Hererra, for their useful guidance, insightful comments, and considerable encouragement to complete this thesis. They have guided me to pursue important problems that will have a practical impact and were always available to guide me whenever I approached them. This work would not have been completed without their encouragement and patience. Also, I would like to express my gratitude to Dr. Hatem A. Rashwan and Dr. Mohamed Abdel-Nasser for their support and assistance. They consistently allowed this thesis to be my work but steered me in the right direction whenever they thought I needed it. Without their precious support, it would not be possible to conduct this research. Also, I wish to acknowledge the assistance and support of Dr. Eleni Mangina from the University College Dublin (UCD), School of Computer Science, who provided me an opportunity to join their team as visiting research student. For three months. I would like to thank all the people at the UCD for giving me this opportunity to stay with them.

Furthermore, I would like to acknowledge all the funding institutions that have supported this research: including the Martí-Franquès Research Fellowship Programme - Doctoral Grant provided by Universitat Rovira i Virgili with the support of Fundació Catalunya La pedrera, Institut de Recerca i Tecnologia Agroalimentària (IRTA), and Spanish project: DPI2016-77415-R and the program of Beatriu de Pinós (B.P. 2016): 2016 BP 00063. Finally, I also thank my friends Pezhman, Hande, Parya, Elena, and Julian, whom I can always count on, for sharing my woes and happiness.

Finally, I must express my profound gratitude to my parents and brother for providing me with unfailing support and continuous encouragement throughout my years of study and throughout the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Author

Nasibeh Saffari.





## Contents

Abstract.....	v
Resumen (Spanish).....	vii
Resum (Catalan).....	x
Acknowledgments.....	xiii
Chapter 1. Introduction.....	11
1.1. Background.....	12
1.2. Objectives and scope of the research.....	14
1.5. Scientific dissemination.....	16
Chapter 2. Feature Learning for Breast Tumor Classification Using Bio-Inspired Optimization Algorithms.....	18
2.1. Introduction.....	19
2.2. The proposed feature learning method.....	22
2.2.1. Whale optimization algorithm (WOA).....	24
2.2.2. The fitness functions.....	26
2.3 Experimental results and discussion.....	28
2.4 Conclusion and future work.....	31
Chapter 3. On Improving Breast Density Segmentation using Conditional Generative Adversarial Networks.....	32
3.1. Introduction.....	33
3.2. Proposed Methods.....	37
3.3. Experimental results.....	40

3.3.1. Database.....	40
3.3.2. Pre-processing .....	40
3.3.3. Result and discussion.....	41
3.4. Conclusion.....	43
Chapter 4. Fully Automated Breast Density_Segmentation and Classification_using Deep Learning.....	45
4.1. Introduction .....	46
4.2. Background Study .....	49
4.2.1. Traditional Computer Vision Methods.....	49
4.2.2. Deep Learning-Based Methods .....	52
4.3. Methodology .....	55
4.3.1. Preprocessing.....	56
4.3.2. Breast Density Segmentation.....	57
4.3.3. Breast Density Classification.....	63
4.4. Experimental Results.....	65
4.4.1. Dataset .....	65
4.4.2. Implementation Details.....	67
4.4.3. Breast Density Segmentation.....	68
4.4.4. Breast Density Classification.....	72
4.5. Conclusions .....	76
Chapter 5. Analyzing the temporal mammographicdensity changes after treatment..	77
5.1. Introduction .....	79

5.2. Methodology .....	83
Implementation details .....	83
5.2.1. Global Breast density follow up analyzing.....	84
5.2.2. Local density follows up analyzing .....	87
5.2.3. Breast density changes overall index.....	89
5.3. Experimental results .....	89
5.3.2. Result of breast density segmentation analysis .....	90
5.4. Conclusions .....	92
Chapter 6. Concluding remarks .....	93
6.1. Thesis highlights .....	94
6.2. Future research lines.....	97

## **Nomenclature**

CAD	Computer-Aided Design
BDS	Breast Density Segmentation
BDP	Breast Density Percentage
BDC	Breast Density Classification
cGAN	Conditional Generative Adversarial Network
DL	Deep Learning
CNN	Convolutional Neural Network
WHO	World Health Organization
CC	Craniocaudal
MLO	Mediolateral Oblique
MD	Mammographic Density
BI-RADS	Breast Imaging Reporting and Data System
ML	Machine Learning
MIAS	Mammographic Image Analysis Society
ANN	Artificial Neural Network
GA	Genetic Algorithm
HOG	Histogram of Oriented Gradients
AUC	Area Under the Curve
ROC	Receiver Operating Curve

WOA	Whale Optimization Algorithm
MLP	Multi-Layer Perceptron
TPR	True Positive Rate
TNR	True Negative Rate
FPR	False Positive Rate
FNR	False Negative Rate
BUS	Breast Ultrasound
P.D.	Percent Density
FCM	Fuzzy c-Means
SVM	Support Vector Machine
GLCM	Gray Level Co-occurrence Matrix
BCE	Binary Cross Entropy
FCN	Fully Convolutional Network
PBC	Primary Biliary Cholangitis
LIBRA	Laboratory for Individualized Breast Radiodensity Assessment
FFDM	Full Field Digital Mammography
CV	Cross Validation
MT-CSAE	Mammographic Texture by Sparse Convolutional Autoencoder
MT	Mammographic Texture
DCNN	Deep Convolutional Neural Network
MSE	Mean Square Error
SSIM	Structural Similarity Index Measure
FC	Fully Connected
2-D	2-dimensional

DSC	Dice Similarity Coefficient
JI	Jaccard Index
BCR	Breast Cancer Recurrence
MRI	Magnetic Resonance Imaging
ER	Estrogen Receptor
PR	Progesterone Receptor
HER2	Human Epidermal Growth Factor Receptor 2

Figure 1. 1 BIRADS density classification.....	12
Figure 1. 2 Scheme of the PhD thesis_.....	15
Figure 2. 1 The proposed feature learning method. To explanation, in this figure, $P_r$ is set to 3 and $P_c$ is set to 2 .....	23
Figure 2. 2 Examples of breast tumors in BUS images (a) benign and (b) malignant.....	29
Figure 2. 3 Examples of breast tumors in mammograms (a) benign and (b) malignant.....	31
Figure 3. 1 Breast density tissues classification in INbreast mammograms.....	36
Figure 3. 2 Conditional Generative Adversarial Network framework.....	39
Figure 3. 3. U-Net architecture .....	40
Figure 3.4. Fully convolutional neural network architecture (FCN-8) .....	43
Figure 3. 5 Examples of the breast density segmentation: (col 1) breast mammogram images, (col 2) ground truth mask, (col 3) generated masks with our proposed method cGAN-U-net, (col 4) FCN 8, (col 5) FCN 32 and (col 6) VGG-SegNet.....	45
Figure 4. 1 Prevalence, relative risks of developing breast cancer based on Four classes of the BI-RADS density standard (i.e., fatty, scattered fibroglandular density, heterogeneously dense and extremely dense) .....	48
Figure 4. 2 An overview of proposed framework.....	56
Figure 4. 3 (a) Original images. (b) After removing pectoral muscle.....	58
Figure 4. 4 cGAN framework for segmentation.....	59
Figure 4. 5 cGAN framework for breast density segmentation.....	63
Figure 4. 6 Proposed CNN architecture used for breast density classification (second technique) .....	65

Figure 4. 7 Breast density segmentation result of three models with the INbreast dataset....	72
Figure 5. 1 Methodology of breast cancer risk based on temporal breast density analyses...	86
Figure 5. 2 Analysing the longitudinal of Whole breast density in three years (V1 is MLO-R, V2 is MLO-L, V3 is CC-R, V4 is CC-L) .....	88
Figure 5. 3 Analysing the longitudinal of mass region density in three years (W1 is MLO-R, W2 is MLO-L, W3 is CC-R, W4 is CC-L) .....	91
Figure 5. 4 sample of temporal breast density segmentation in three years .....	93
Figure 5. 5 sample of global (a) and local (b) breast density analysing for one non-risky patient with two years follow up.....	94



Table 2. 1 The search range of the four parameters.....	30
Table 2. 2 Results with the BUS database.....	30
Table 2. 3 Results with the Mini-MIAS database.....	32
Table 3. 1 Recall, Specificity, FPR, FNR, PBC, Precision and F-score with the cGAN-Unet, FCN8, FCN32 and VGG-SegNet methods evaluated on the testing set of the INBreast dataset. The best results are marked in a bold text.....	44
Table 4. 1 Distribution of densities across imbalanced INbreast dataset (before augmentation) .....	67
Table 4. 2 Accuracy, DSC, and JI with the cGAN-Unet, cGAN-Unet-SSIM-loss, and cGAN-Unet-dice-loss evaluated on the testing set of the INBreast dataset for breast density segmentation (C1 = Class1, C2 = Class2, C3 = Class3, C4 = Class4) .....	70
Table 4. 3 Summary of Sensitivity, Specificity, FPR, FNR, Precision, and DSC with the cGAN-Unet-MSE-loss, FCN8, FCN32, and VGG-SegNet methods evaluated on the testing set of the INBreast dataset for breast density segmentation.....	73
Table 4. 4 Confusion matrix of breast density estimated based on thresholding rules.....	74
Table 4. 5 Distribution of densities across balanced INbreast dataset (after augmentation) ...	75
Table 4. 6 Accuracy, Precision, Sensitivity, and Specificity of the CNN-based classification method on imbalanced and balanced datasets with two different sizes of input images: 64×64 and 128×128.....	75
Table 4. 7 Confusion matrix of breast density by CNN-based classification method.....	77
Table 4. 8 Accuracy of different breast density classification in terms of their best results....	77
Table 5.1 Mean and standard deviation of different biomarkers (mean ± S.D.) .....	95

Table 5. 1 One-way ANOVA analysis for breast density and different breast cancer parameters before surgery, first control after surgery (C1) and second control after surgery (C2) ..... 95

Table 5. 2 Correlation coefficient before surgery, control 1 and control 2 of breast density with different parameters of breast cancer..... 96

# Introduction

*This first chapter highlights the details about the challenges to finding the contrast between the fibroglandular density and mass tissue and the motivation of analyzing the breast density and explains the significant objectives of this research and how we solve it. Also, it presents the thesis structure and some of the leading publications concluded from the thesis.*

## 1.1. Background

According to World Health Organization (WHO), doctors cannot diagnose 10-30% of breast masses(Chen, Wang, & Yang, 2006; Figueredo & Wolf, 2009) as both mass and fibroglandular tissue occur white on imaging and may be mistaken, which leads to concealing the presence of breast cancer. Generally, a mammogram's Computer-aided diagnosis (CAD) is generated to increase the sensitivity and accuracy and assist radiologists in diagnosing disease by providing processing results to the doctors as a second opinion(Tai, Chen, & Tsai, 2014). Nowadays, computer technologies and image processing techniques have been available for radiologists to help them achieve automated mammogram computer-aided detection (CAD) and give them more facilities. Regardless, most medical practitioners believe that early detection of breast cancer will increase the patient's chance to survive and save many thousands of lives each year.

The World Health Organization has reported in the World Cancer Report that mammographic density (M.D.) is a "well-characterized" risk factor of breast cancer(Stewart & Wild, 2015) and a significant determinant of screening sensitivity. In addition, breast density is an accepted risk factor for breast cancer through increasing exposure over the lifetime. Many studies compare women with extremely dense breast tissue with more than 75% density to women with a percent density of <5%, with the former group having four to six times higher risk for breast cancer(Harvey & Bovbjerg, 2004). The risk of second breast cancers between women with highly dense breasts has also been observed three times(Habel et al., 2004). Breast density is the absolute amount or percentage of fibroglandular tissue in the breast, which can be observed white and bright on a mammogram and is also known as mammographic density. The female breast consists

of fat and fibroglandular tissue. Fat is radiographically lucent and appears dark on a mammogram.

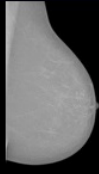
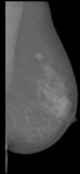
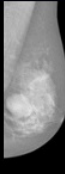
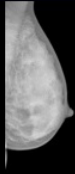
Example				
Category	BI_RADS I Fatty breast	BI_RADS II Average density	BI_RADS III Heterogeneously dense	BI_RADS IV Extremely dense

Figure 1. 1 BIRADS density classification.

In contrast, fibroglandular tissue is radiographically dense and appears white on a mammogram. Therefore, fibroglandular tissue is also known as dense tissue, and fat is often referred to as non-dense tissue. Breast density varies among women, and in most but not all women, it declines with age, especially around menopause, also can be affected by factors like changing hormone levels (e.g., use of estrogen, place in the menstrual cycle, etc.), genetic factors, parity, use of tamoxifen, weight.

The look, size, or feel of a breast does not provide any information about breast density. Breast density is a radiological finding measured by mammography (either by a radiologist reviewing mammographic images or via computer-aided assessment). While there are different ways to measure breast density, radiologists usually classify mammographic images via computer-aided assessment. Figure 1.1 shows the differences in density using Breast Imaging Reporting and Data System (BI-RADS) classification.

In addition, Deep Learning (DL) is a new trend in Machine Learning (ML), inspired by simulating the structure and function of the brain called artificial neural networks. It has opened an opportunity to make better and faster learning algorithms. The core of deep learning is training the large neural network using fast computers with a large amount of

data, bigger models, and more computation to achieve better performance. It is fantastic at supervised learning.

Deep learning techniques are increasingly used for breast density analysis to automate the features extraction of mammograms at multiple levels of abstraction and evidence of superior performance, breast density segmentation, and breast density classification. Deep learning networks, such as Convolutional Neural Networks (CNN), can automatically learn features from raw images directly and accurately represent objects at different scales and orientations. A convolutional neural network (CNN) is one of the most popular algorithms for deep learning.

In short, our motivation is to analyze the breast density using deep learning methods to provide the more accurate breast density CAD to help doctors find the breast density easier and faster.

## **1.2. Objectives and scope of the research**

The main objective of this thesis is to create an advanced CAD system for analyzing the breast density, such as segmentation, classification, and analyzing the temporal mammograms to find the breast density changes after treatment or surgery and in the following predicting the risky patients. The significant objective of this thesis detailed are as follows:

1. Due to the association of breast density and breast tumor types, a new method based on a whale optimization algorithm was presented to discriminate between malignant and benign tumors by combining the traditional feature extraction and deep learning methods to automatically learn local descriptors from the input images without handcrafting the feature extraction methods. This chapter is taken from a paper presented

- in the ITCE Conference entitled feature learning for breast tumor classification using bio-inspired optimization algorithms.
2. A novel methodology for automating breast density segmentation to enhance the accuracy of breast density segmentation using the cGAN network with MSE loss function developed. This chapter is taken from a paper presented at the CCIA Conference entitled “Improving Breast Density Segmentation using Conditional Generative Adversarial Networks.”
  3. The breast density estimation by our multiclass CNN architecture correlated well with BI-RADS density ratings (BI-RADS I, BI-RADS II, BI-RADS III, BI-RADS IV ) using the binary mask segmented in the previous stage (cGAN output) by radiologists. This chapter is taken from a "Diagnostics" Journal paper entitled “Fully Automated Breast Density Segmentation and Classification Using Deep Learning.”
  4. Another objective was to analyze the temporal breast density changes after surgery. In this regard, the combination of the whole breast density (global region) and mass region density (local area) were analyzed based on the correlation between the output of our automated algorithm and the radiologist's presented breast density measures to find the risky patients after surgery.

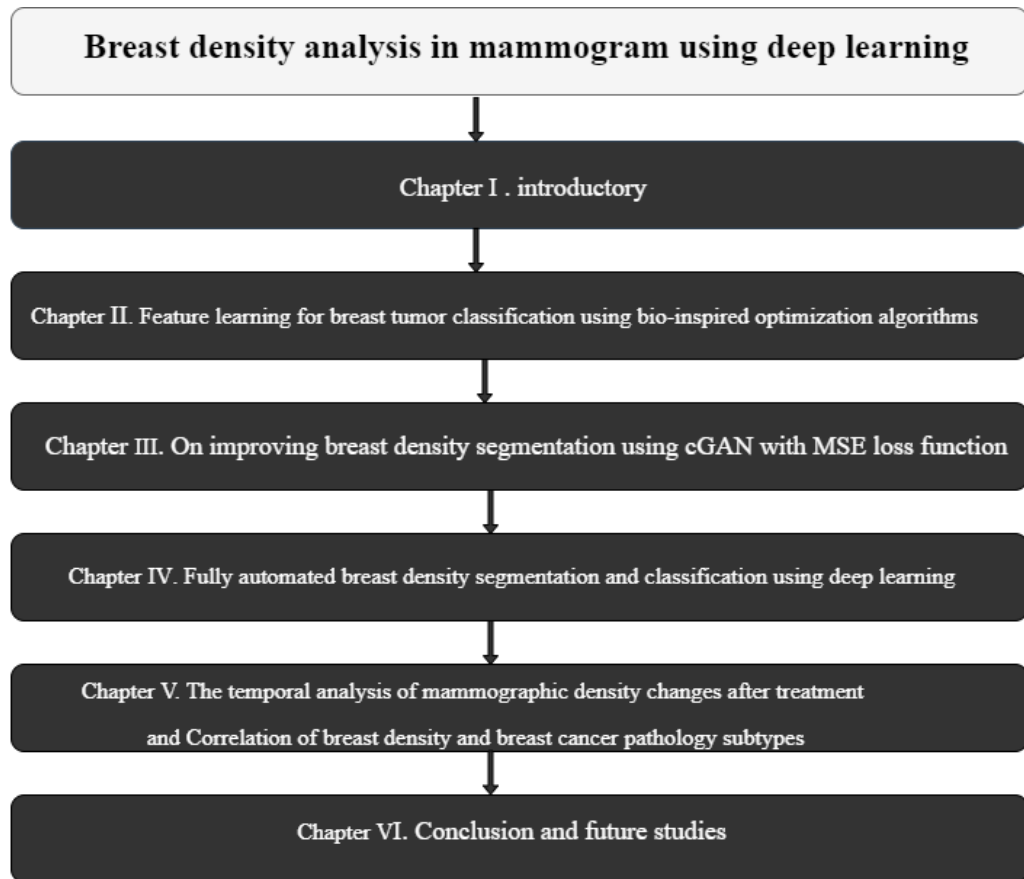


Figure 1. 2 Scheme of the Ph.D. thesis.

## 1.5. Scientific dissemination

### 1- Feature learning for breast tumour classification using bio-inspired optimization

**algorithms**, Conference: 20<sup>th</sup> International Conference of the Catalan Association for Artificial Intelligence (CCIA 2017). *Authors: Adel Saleh, Mohamed Abdel-Nasser, Antonio Moreno, Nasibeh Saffari and Domènec Puig.*

### 2- Data visual embedding for image classification

, Conference: 2018 International conference on innovation trends in computer engineering (ITCE). *Authors: Adel Saleh, Mohamed Abdel-Nasser, Md. Mostafa Kamal Saker, Vivek Kumar Singh, Saddam Abdulwahab, Nasibeh Saffari, Miguel Angel Garcia, and Domènec Puig.*



- 3- On improving breast density segmentation using a conditional generative adversarial network**, 21<sup>st</sup> International Conference of the Catalan Association for artificial intelligence, CCIA 2018, Roses, Spain.). *Authors: Nasibeh Saffari, Hatem Rashwan, Blas Herrera, Santiago Romani, Meritxell Arenas and Domènec Puig.*
- 4- Fully automated breast density segmentation and classification using deep learning," Diagnostics Journal paper.** ( Nov 2020). *Authors: Nasibeh Saffari, Hatem Rashwan, Mohamed Abdel-Nasser, Vivek Kumar Singh, Eleni Mangina, Meritxell Arenas Blas Herrera, and Domènec Puig.*
- 5- Temporal analysis of mammographic density changes after treatment, on going Journal paper.** *Authors: Nasibeh Saffari, Hatem Rashwan, Mohamed Abdel-Nasser, Meritxell Arenas Blas Herrera, and Domènec Puig.*

# Feature Learning for Breast Tumor Classification Using Bio-Inspired Optimization Algorithms

*Breast cancer is one of the most dangerous diseases that threaten women. Computer-aided diagnosis systems can be used to analyze breast images to detect breast tumors early. This chapter proposes a new method for feature learning from breast cancer images. The learned features are used to discriminate between malignant and benign tumors. We marry the traditional feature extraction methods with deep learning approaches because we automatically learn local descriptors from the input images without handcrafting the feature extraction methods. The proposed technique uses a bio-inspired optimization method, called the whale optimization algorithm, to learn local descriptors from the input images. The learned features are tuned to the input images and can describe the local/global arrangement of pixels in them. Unlike the usual deep learning approaches, small datasets can train the proposed method. We used two breast cancer datasets to validate the proposed feature learning method: mammographic and ultrasound images for benign and malignant cases. The experimental results demonstrate that the proposed method gives good classification results than the state-of-the-art ones.*

## 2.1. Introduction

Breast cancer is one of the most dangerous diseases for women. Mammography is the most common method for the early detection of breast cancer, although ultrasound images are the common method to detect breast cancer in dense breasts (young women). Radiologists investigate mammograms to identify abnormalities. Computer-aided diagnosis (CAD) systems typically use machine learning methods to detect tumors in breast images. Such techniques need to work with discriminant and defining features to classify breast images into multiple classes. Several methods have been proposed to classify breast tumors and attain good classification accuracy. However, it is still an open research area due to the inherent challenges in breast image representation and classification (e.g., the noise and artifacts in mammograms and ultrasound images).

Several methods have been proposed to analyze breast tumors using mammograms in the literature. Jadoon et al. proposed two methods (convolutional neural network-discrete wavelet and convolutional neural network-curvelet transform) to classify breast tissues into normal, benign, or malignant (Abdel-Nasser, Saleh, Moreno, Tabalvandani, & Puig, 2017). The mini-mammographic Image Analysis Society (MIAS) database achieved a sensitivity of 0.888 and a specificity of 0.801. The authors of (Setiawan, Elysia, Wesley, & Purnama, 2015) used Law's text pure energy measures texture analysis method with an artificial neural network (ANN) to classify mammograms into normal/abnormal and benign/malignant images. They achieved an accuracy of 53:06% with the mini-MIAS database. Two automated methods were presented to analyze breast tumors in mammograms (Rouhi, Jafari, Kasaei, & Keshavarzian, 2015). In the first method, image segmentation was done using an automated region growing algorithm whose threshold was obtained by a trained ANN. In the second method, image segmentation was performed

by a cellular neural network whose parameters were determined by a genetic algorithm (G.A.). Intensity, textural, and shape features were extracted from segmented tumors. A GA algorithm was used to select an appropriate set of features. Finally, the ANN classifier was used to classify the mammograms as benign or malignant. In (Abdel-Nasser, Melendez, Moreno, Omer, & Puig, 2017), super-resolution was used to improve the performance of texture analysis methods when classifying breast tumors in ultrasound images.

Indeed, texture analysis methods are widely used to characterize breast tumors in mammograms and ultrasound images. Texture produces information about the spatial arrangement of the intensities in the suspicious regions. The effect of pixel resolution, integration scale, preprocessing, and feature normalization methods on the performance of texture analysis methods for breast tumor classification was studied in (Abdel-Nasser, Melendez, Moreno, & Puig, 2016). Five handcrafted texture methods were analyzed: local binary pattern, local directional number, histogram of oriented gradients (HOG), Haralicks features, and Gabor filters. They concluded that the factors mentioned above affect the performance of texture methods, so the best combination of the factors should be determined to obtain the best performance with each texture analysis method. Several studies have used deep-learning methods (Arevalo, González, Ramos-Pollán, Oliveira, & Guevara Lopez, 2016; Becker et al., 2017; Dhungel, Carneiro, & Bradley, 2017; Jadoon, Zhang, Haq, Butt, & Jadoon, 2017; J. Wang et al., 2016), especially deep convolutional neural networks (CNN). The authors of (Litjens et al., 1998) reviewed the major deep learning concepts pertinent to medical image analysis. The authors of (Dhungel et al., 2017) proposed a method for segmenting and classifying breast masses in mammograms with minimal user intervention. The authors of (J. Wang et al., 2016) concluded that deep learning based on large datasets was superior to standard methods for the discrimination

of microcalcifications in mammograms. Arevalo et al. (Arevalo et al., 2016) showed that the combination of deep learned and handcrafted features produced the best descriptor for mass classification, obtaining an AUC of 0.826 (the area under the curve (AUC) of the receiver operating curve (ROC)).

Although deep learning techniques remove the burden of handcrafting the texture features, applying them to medical image analysis has limitations. The main limitation of using deep learning methods with medical image analysis is that they require thousands of images to train their models. In breast cancer, only small datasets are available (tens to hundreds of images); thus, it is difficult to use deep learning approaches to extract features from breast cancer images. Some researchers use transfer learning approaches to overcome the problem of small-sized datasets, where they use models trained on nature images, e.g., AlexNet (Krizhevsky, Sutskever, & Hinton, 2012), and then they fine-tune the weights of the models using the medical images. However, the transfer learning approach does not give accurate results with all medical datasets because the medical images have different characteristics than the nature images.

Moreover, medical images Usually contain noise, which significantly degrades the performance of deep learning models. Another limitation is that the deep CNN approach is not a fully automatic feature learning method because the user should design the network and set the values of many critical parameters, such as the number of convolutional layers, the number of filters per layer, the number of the fully connected layers, and whether to use a pooling layer or not.

This chapter proposes a new method for feature learning from breast cancer images. The learned features are used to classify breast tumors into malignant or benign. In this method, we marry the traditional handcrafted feature extraction mechanisms with

deep learning approaches, where we automatically learn descriptors from breast images. The proposed technique uses a bio-inspired whale optimization algorithm (WOA) to learn the local descriptors from breast images. Indeed, we use the WOA algorithm because it gives accurate optimization results when solving several mathematical and structural problems (Mirjalili & Lewis, 2016). The learned features are tuned to the input set of images and characterize the local/global information. Unlike the deep learning approaches, small-sized datasets can be used to train our method.

The rest of this chapter is organized as follows. In Section 2, we explain the proposed feature learning method. Then, section 3 presents and discusses the experimental results. Finally, section 4 concludes the paper and suggests some points of future work.

## 2.2. The proposed feature learning method

Figure. 2.1 shows the steps of the proposed method. First, each input image is divided into  $N_p = P_r \times P_c$  Non-overlapped patches, where  $P_r$  is the number of patches per row and  $P_c$  is the number of patches per column. Next, the method extracts local features from each patch using HOG descriptors (Mirjalili & Lewis, 2016), and each patch is divided into  $B$  blocks containing  $C$  cells. The occurrences of edge orientations are computed, and a weighted histogram is computed for each block. The frequencies in the histograms are normalized in the interval  $[0,1]$  to compensate for changes in illumination. Next, the histograms of all blocks are concatenated to form a descriptor  $F_p$  for the given patch. Finally, the descriptors of all patches are concatenated to form a descriptor  $F_{img}$  for the whole input image as follows:  $F_{img} = \cup_{i=1}^{N_p} F_p(i)$ , where  $\cup$  refers to feature concatenation.

The WOA determines the optimal values of  $P_r$ ,  $P_c$ ,  $B$  and  $C$  produce descriptors adapted to describe the characteristics of the breast tissues in the input images. Below, we show the

population of the WOA, where each row in the population represents a search agent. The algorithm uses  $N$  search agents to find the optimal solution, where each search agent  $SA_j$  contains candidate values for the four parameters  $P_r$ ,  $P_c$ ,  $B$ , and  $C$  ( $SA_j = [P_{r_i}; P_{c_i}; B_i; C_i]$ ,  $i = 1, 2, \dots, N$ ).

$$\text{Population} = \begin{bmatrix} P_{r_1} & P_{c_1} & B_1 & C_1 \\ P_{r_2} & P_{c_2} & B_2 & C_2 \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ P_{r_N} & P_{c_N} & B_N & C_N \end{bmatrix}$$

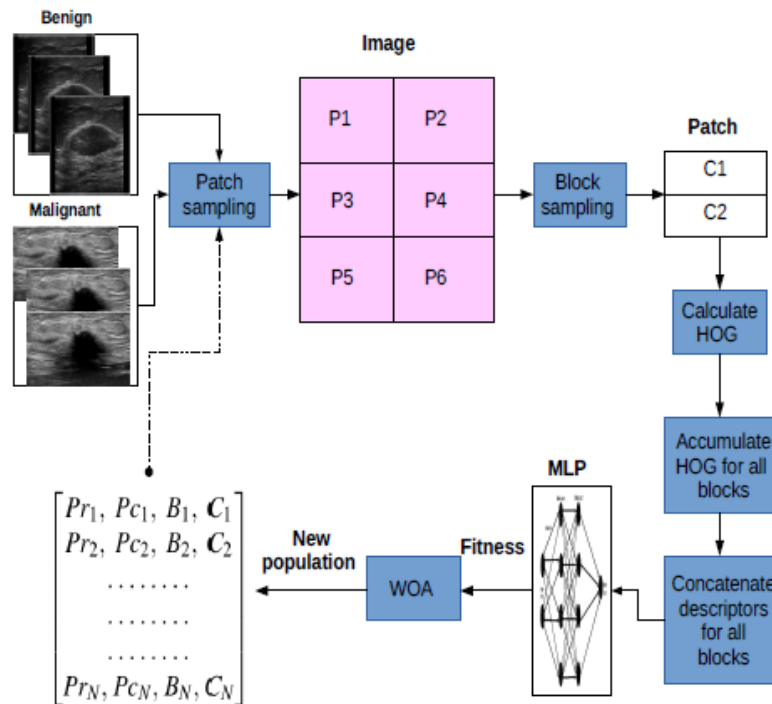


Figure 2. 1 The proposed feature learning method. To explanation, in this figure,  $P_r$  is set to 3, and  $P_c$  is set to 2

The proposed method receives  $T_b$  images of benign tumors and  $T_m$  of malignant tumors. For a given search agent  $SA_j = [P_{r_i}; P_{c_i}; B_i; C_i]$ , the proposed method computes the features using the values of the four parameters for the benign ( $Feat_b$ ) and the malignant images ( $Feat_m$ ). The learned features  $Feat_b$  and  $Feat_m$  are fed into a multi-layer perceptron (MLP) to classify the images into benign or malignant. Then, we use the AUC of ROC to calculate the fitness function of WOA using the following formula:  $fitness = 1 - AUC$ . In the subsection below, we explain the steps of WOA and how we calculate the fitness function.

### 2.2.1. Whale optimization algorithm (WOA)

Whales are considered brilliant animals. They have spindle cells in their brains responsible for their social behavior, emotions, and judgment. Humpback whales, which are giant whales, have a unique hunting method. They prefer to hunt a krill group or small fishes close to the sea's surface. First, whales make bubbles along a circle or '9'-shaped path to hunt prey (bubble-net feeding behavior). Then, the whales determine the location of prey and encircle them. Since the exact position of prey is unknown, WOA assumes that the current candidate solution is the position of the prey, or it is a position that is close to the optimum one. The main steps of hunting the prey are encircling, the spiral bubble-net feeding maneuver (exploitation stage), and the search for the prey (exploring stage). In the following subsections, we explain the mathematical model of the three steps.



### 2.2.1.1. The model of encircling the prey

After WOA defines the best search agent, the other agents update their positions towards the position of the best search agent. This behavior can be formulated as follows (Mirjalili & Lewis, 2016):

$$D = |\vec{C} \cdot \vec{X}^*(t) - \vec{X}(t)| \quad (2.1)$$

$$X^*(\vec{t} + 1) = \vec{X}^*(t) - \vec{A} \cdot \vec{D} \quad (2.2)$$

In this expression,  $t$  indicates the current iteration,  $\vec{A}$  and  $\vec{C}$  are coefficient vectors, The position vector of the best solution is  $X^*(t + 1)$ ,  $\vec{X}(t)$  is the position vector, is an element-wise multiplication, and  $jj$  is the absolute value. WOA updates the value of  $X^*(t+1)$  when it finds a better solution.  $\vec{A}$  and  $\vec{C}$  can be determined using the following expressions:  $\vec{A} = 2 \cdot \vec{a} \cdot \vec{r} - \vec{a}$ ,  $\vec{C} = 2 \cdot \vec{r}$ . Where  $\vec{a}$  is linearly decreased from 2 to 0 over the iterations of WOA, and  $\vec{r}$  is a random vector in the interval  $[0,1]$ .

### 2.2.1.2. The model of spiral bubble-net feeding

There are two approaches to simulate the bubble-net feeding behavior: the shrinking encircling mechanism and the spiral updating position.

**Shrinking encircling mechanism:** We can achieve this behavior by decreasing

The value of  $\vec{a}$  over the iterations of WOA.  $\vec{A}$  is a random value in the interval  $[-a, a]$  The new position of the search agent can be determined by setting random values for  $\vec{A}$  in the interval  $[-1,1]$ .

**Spiral updating position:** Assume that the whale's location is  $(X; Y)$  and the location of the prey is  $(X^*; Y^*)$ . This approach determines the distance between them and then uses a spiral equation to simulate the helix-shaped movement of humpback whales as follows:

$$\vec{X}(t + 1) = \vec{D}^i \cdot e^{bl} \cdot \cos(2\pi l) + \vec{X}^*(t) \quad (2.3)$$

In this equation,  $b$  is a constant determining the shape of the logarithmic spiral,  $l$  is a random value in the interval  $[-1,1]$ , and  $\vec{D}^i$  is the distance between the humpback whale and the prey can be calculated as follows:  
 $\vec{D}^i = |\vec{X}^*(t) - \vec{X}(t)|$ .

The humpback whales swim around the prey within a shrinking circle or spiral-shaped path. To simulate this behavior, WOA assumes a probability of 50% to use one of the two mechanisms (circular or spiral).

### 2.2.1.3. The model of the search for the prey

This model also varies the value of  $\vec{A}$  To find the prey. Unlike the exploitation phase, in this model, WOA updates the position of the search agent randomly from the population instead of using the position of the best search agent. This model can be expressed as follows:

$$D = |C \cdot \overrightarrow{X_{rand}^*}(t) - \vec{X}(t)| \quad (2.4)$$

$$X(\vec{t} + 1) = \overrightarrow{X_{rand}^*}(t) - \vec{A} \cdot \vec{D} \quad (2.5)$$

Where  $|\vec{A}| < 1$  and  $\overrightarrow{X_{rand}^*}$  is a random position vector.

### 2.2.2. The fitness functions

In each iteration of the proposed feature learning method, we calculate the texture features using the values of each search agent, and then we input the features of the benign and malignant classes into an MLP classifier to calculate the fitness corresponding to this agent. Finally, we use the AUC of the ROC curve to calculate the fitness function as follows:

$$fitness = 1 - AUC \quad (2.6)$$

The following terms are defined to compute the AUC values:

\_ **True Positive (T.P.):** malignant tumor classified as malignant.

\_ **True Negative (T.N.):** benign tumor classified as benign.

\_ **False Positive (F.P.):** benign tumor classified as malignant.

\_ **False Negative (F.N.):** malignant instance classified as negative.

The ROC analysis is used to avoid selecting a single threshold for classification.

To determine the ROC curve of the MLP model, we scan through all possible thresholds and record the true positive rate  $\frac{TP}{TP+FN}$  and the false positive rate  $\frac{FP}{FP+TN}$  (Dhungel et al., 2017).

The MLP classifier is a network of simple neurons called perceptrons. Each perceptron determines a single output from numerical inputs by calculating a weight for each input and then combining them. The resulting value is then input into a non-linear activation function. The operation of MLP can be expressed as follows:

$$y = \phi \left( \sum_{i=1}^n w_i x_i + b \right) \quad (2.7)$$

where  $y$  is the perceptron's output,  $x$  is a vector containing the inputs,  $w$  is a vector containing the weights,  $b$  is the bias and  $\phi$  activation function. The hyperbolic tangent  $\phi(z) = \tanh(z)$  activation function is used in this chapter. We used the MATLAB function `divider` to randomly divide the dataset into training, validation, and testing sets. We set the training, validation, and testing ratios to 0.7, 0.15, and 0.15, respectively.

## 2.3 Experimental results and discussion

To validate the proposed method, we present two sets of experiments in this section:

\_ We used the proposed method to discriminate between benign and malignant tumors in ultrasound images. We used a breast ultrasound (BUS) database created by the Engineering Department of Cambridge University. The complete database is available on Peipa website<sup>1</sup>.

\_ We used the proposed method to discriminate between benign and malignant tumors in mammograms. We used the mini-MIAS database in this experiment. The BUS database consists of 31 malignant and 28 benign BUS images. This dataset is part of a clinical database of ultrasonic radio frequency strain imaging data. In Figure. 2.2 we present examples of benign and malignant tumors in BUS images.

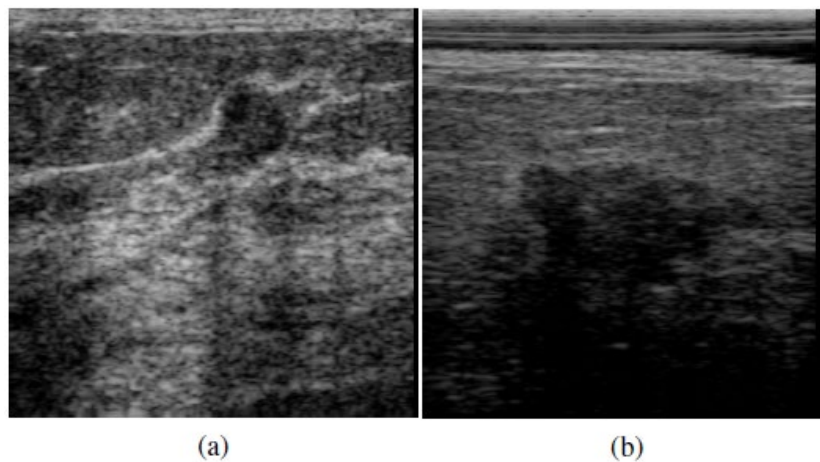


Figure 2. 2 Examples of breast tumors in BUS images (a) benign and (b) malignant.

The mini-MIAS database consists of 322 mammograms of 161 cases. It was created from the original MIAS database by down sampling the images from 50mm to

<sup>1</sup> <http://peipa.essex.ac.uk/info/mias.html>

200mm per pixel and clipping/padding to a fixed size of 1024\_1024 pixels. Experienced radiologists prepared the ground truth of the mini-MIAS database, which was confirmed using a biopsy. This chapter used 60 regions of interest (ROIs) containing benign masses and 49 ROIs containing malignant tumors. In Figure. 2.3, we present examples of benign and malignant tumors in mammographic images.

**Table 2. 1 The search range of the four parameters**

<b>Parameter</b>	<b>Min.</b>	<b>Max.</b>
$P_r$	1	4
$P_c$	1	4
<b>B</b>	1	4
<b>C</b>	1	4

Table 2.1 shows the minimum and maximum values (the search range) of the four parameters, i.e.,  $P_r$ ,  $P_c$ , B, and C. The number of search agents was 10. The maximum number of iterations was 100.

**Table 2. 2 Results with the BUS database.**

<b>Methods</b>	<b>AUC</b>	<b>Best parameter [<math>P_r</math>, <math>P_c</math>, B, C]</b>
<b>Proposed</b>	0.94	[3,3,2,1]
<b>HOG</b>	0.828	-

Table 2.2 shows the results of the proposed method with the BUS images. It obtains an AUC of 0.94, which corresponds to the optimal values of  $P_r$ ,  $P_c$ , B, and C. We also compare the proposed method with the results of the HOG descriptor in the related study of (Abdel-Nasser, Melendez, et al., 2017), which used the same dataset and achieved an AUC of 0.828.



Figure 2. 3 Examples of breast tumors in mammograms (a) benign and (b) malignant

Table 2.3 shows that we achieved an AUC of 0.83 with the mini-MIAS database, while the study in (Abdel-Nasser, Melendez, et al., 2016) achieved an AUC of 0.78 with the local binary pattern (LBP) and an AUC of 0.74 with the HOG method. The proposed method obtains good classification results, and it is on par with the related methods. Our method provides a good description of the gradients of the input images, and it is less sensitive to noise. Therefore, it gave good results with the BUS images, which mainly suffer from the speckle noise. Although mammograms suffer from several artifacts, blurring and noise, our method also gave a good AUC value.

The study of (Abdel-Nasser, Melendez, et al., 2016) concluded that we should find the best combination of four factors (pixel resolution, integration scale, preprocessing, and feature normalization methods) to obtain the best performance with each texture analysis method. Unlike traditional texture analysis methods (e.g., LBP, Gabor features and HOG), the proposed technique does not need handcrafting or selecting the best combination of the abovementioned factors.

Table 2. 3 Results with the Mini-MIAS database.

Methods	AUC	Best parameters $P_r, P_c, B, C$
<b>Proposed</b>	0.83	[2,2,1,2]
<b>LBP</b>	0.78	-
<b>HOG</b>	0.74	-

Unlike the deep learning approaches (e.g., the deep CNN), training the proposed method does not require thousands of images, and we can use tens of images in the training phase. In general, it is challenging to find big medical image databases because labeling thousands of medical images is costly and time-consuming. Indeed, our method is very suitable for tumor classification in breast images as we have relatively small datasets.

## 2.4 Conclusion and future work

In this chapter, we have proposed a new method for feature learning from breast cancer images. We used the learned features to classify breast tumors into malignant or benign. In the proposed method, we marry the traditional texture analysis methods with deep learning approaches, where we automatically learn local descriptors from the input images without handcrafting them. Our method uses the whale optimization algorithm (a bio-inspired optimization method) to learn local descriptors from the input images. The learned features are tuned to the ultrasound and mammographic images and characterize the spatial arrangement of the intensities in the images. Unlike the deep learning approaches, a small dataset can be used to train our method. We used ultrasound and mammographic images datasets with benign and malignant cases to assess the proposed method. It gave an AUC of 0.94 with the ultrasound images and an AUC of 0.83 with mammographic images. We also compared it with related methods, finding that the proposed method is on par with the other methods. To further improve the results, the future work will focus on using other texture analysis methods such as Gabor wavelets and empirical wavelets as a basis descriptor (instead of using the oriented gradients). We will also assess the value of adding more layers to the proposed feature learning method, such as pooling,

rectification, and hashing layers. In addition, we will use more medical image datasets to demonstrate the effectiveness of the proposed method.

## CHAPTER 3

# On Improving Breast Density Segmentation using Conditional Generative Adversarial Networks

*Breast density (B.D.) is crucial to follow-up the breast cancer relapse in mammograms. Accurate B.D. Estimation with visual assessment is still a challenge due to faint contrast and significant variations in background fatty tissue. The critical key of accurate breast density estimation is to properly detect the dense tissues in a mammographic image. Thus, this chapter proposes an automatic deep breast density segmentation using conditional Generative Adversarial Networks (cGAN) that consist of two successive deep networks: generator and discriminator. The generator learns the mapping from the input mammogram to the output binary mask preserving the border of the dense tissues. In turn, the discriminator learns a loss function to train this mapping by comparing the ground-truth and the predicted mask under observing the input*



*mammogram as a condition. The performance of the proposed model was evaluated on the public INBreast mammographic datasets. The proposed model can segment the dense regions with overall recall, precision, and F-score around 95%, 92%, and 93%, respectively, outperforming the state-of-the-art of breast density segmentation. The proposed model can perform segmentation of more than 40 images with a 512x512 size per second on a recent GPU.*

## **3.1. Introduction**

Breast cancer is the most common cause of death among women, prevalent in both developed and developing countries. However, early detection of breast cancer improves the treatment of the patients and increases their chance of survival. Several risk factors for breast cancer have been recognized, including age, family profile, genetics, and breast density. Since breast masses have a very similar appearance to dense tissues, many studies have indicated that breast density is a strong risk factor for breast cancer (Arнау Oliver, Jordi Freixenet, Robert Marti, Josep Pont, Elsa Pérez, Erika R. E. Denton, 2008; He, Denton, Stafford, & Zwiggelaar, 2011; Mac Parthaláin, Jensen, Shen, & Zwiggelaar, 2010; Oliver et al., 2015; Sanjay-Gopal et al., 1999). In addition, mammographic density has been associated with the risk of local recurrence after conservative surgery or radiotherapy, as it is a predictor of local recurrence, especially in patients who have not received it (Park, Rembert, Chew, Moore, & Kerlikowske, 2009). Therefore, accurate breast density estimation is essential during the screening procedure because women with dense breasts

tissues have a three-fold to six-fold higher probability of facing the risk of breast cancer (Breast Cancer U.S, 2016; Cancer Research U.K., 2014; Whitman, 2010). Thus, breast density estimation is used to predict the presence of tumors at the early stage to help doctors plan appropriate treatment, either chemotherapy or radiotherapy.

The amount of fibro glandular tissue content in the breast is commonly referred to as breast percent density (P.D. %) (Keller et al., 2012a). According to the research presented by Norman et al.(Linver, 2008), one of the factors that hinder the detection of masses by the specialists is the type of breast density, which can be dense (fibrous) or non-dense (fat). Furthermore, Breast Imaging Reporting and Data System standard (BI\_RADS) (Orel, Kay, Reynolds, & Sullivan, 1999), presented by the American College of Cancer, provides four categories: BI\_RADS I, BI\_RADA II, BI\_RADS III, and BI\_RADS IV, as shown in Figure 3.1.

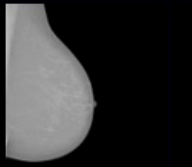
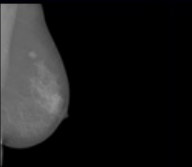
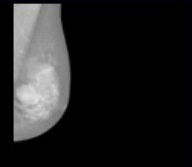

<b>Example</b>				
<b>Category</b>	<b>BI_RADS I</b> Fatty breast	<b>BI_RADS II</b> Average density	<b>BI_RADS III</b> Heterogeneously dense	<b>BI_RADS IV</b> Extremely dense
<b>Range</b>	<b>0-25%</b>	<b>26-50%</b>	<b>51-75%</b>	<b>75-100%</b>

Figure 3. 1 Breast density tissues classification in INbreast mammograms.

X-ray mammography is considered the most popular method utilized by radiologists for screening and early detection. The standard screening mammographic views are craniocaudally (CC) and mediolateral oblique (MLO). The CC mammographic view is captured from the superior view of a horizontally compressed breast. In turn, the

MLO view is captured from the side of a diagonally compressed. Analysis of mammograms is not easily feasible for every case; for this reason, computer techniques to diagnose and classify the breast density have attracted the researcher's attention in the recent decade (Arнау Oliver, Jordi Freixenet, Robert Marti, Josep Pont, Elsa Pérez, Erika R. E. Denton, 2008; Gamdonkar, Tay, Ryder, Brennan, & Mello-Thoms, 2015; Keller et al., 2012a; Oliver et al., 2015; Rampun, Scotney, Morrow, Wang, & Winder, 2018; Sun, Sc, Stone, Sc, & Fishell, 2007). Several approaches for breast density estimation in mammogram analysis have been developed and presented in the literature. These methods can be classified into *handcrafted-based* methods and *deep learning-based* methods. *Handcrafted* methods extract features from mammograms to represent the dense tissues. Traditional machine learning algorithms are used to segment or classify the features. For instance, a fully automated quantification of breast density percent based on adaptive multiclass fuzzy c-means (FCM) clustering and support vector machine (SVM) classification has been proposed by B. M. Keller et al. (Valerie A. McCormack & Dos Santos Silva, 2006). In addition, the LBP descriptor (Zheng et al., 2015), Gabor filters (Gamdonkar et al., 2015), and Gray Level Co-occurrence Matrix (GLCM) features (Petroudi et al., 2015) have been used to extract features from the input mammograms for segmenting and estimating breast density.

In turn, much research has recently been proposed based on *deep learning* systems. Deep learning networks, such as Convolutional Neural Networks (CNN), automatically learn features from raw images directly, and they can adequately represent the object at different scales and orientations. An unsupervised deep learning method based on CNN to learn the characteristics of dense and fatty tissues have been presented by M. Kallenberg (Kallenberg et al., 2016). The input mammogram is divided into a set of sub-images that are classified into dense or fatty regions. Their proposed model strongly correlated to the

manual breast density segmented by expert radiologists. In addition, a deep learning system consisting of four convolutional layers with a maximum pooling has been developed by C. K. Ahn (Kingma & Ba, 2014) to segment the breast density region in mammograms. The convolutional layers are trained as auto-encoders, and their weights and bias terms are fine-tuned using SoftMax regression.

It should be considered that increasing the quality of dense tissue segmentation increases breast density estimation accuracy. The novelty of the current paper is that a deep model has been proposed using conditional Generative Adversarial Network (cGAN) (Ahn, Heo, Jin, & Kim, 2017) to improve the segmentation of the dense breast regions in mammograms. cGAN is a deep learning model that can learn the statistical invariant features (texture, color, etc.) of an input image and then generate nearly synthetic images, which look like the input image. cGAN networks consist of two successive networks: generator and discriminator. First, the generator network learns to map the input image to the target output; in turn, the discriminator learns a loss function to train this mapping by comparing the ground truth and the predicted output. Finally, the cGAN network optimizes a loss function that combines a conventional binary cross-entropy loss with an adversarial term (i.e., discriminator), encouraging the generator to produce outputs that cannot be distinguished from ground-truth ones. The main objective of the present paper is to enhance the accuracy of breast density segmentation using the cGAN network. To the best of the authors' knowledge, this is the first application of the cGAN network to segment dense tissues in mammographic images.

The remaining part of this chapter is structured as follows. First, section 3.2 is devoted to the proposed cGAN model. Then, to demonstrate the proposed work's effectiveness, some experimental results and a comparative study are shown in Section 3. In addition, the conclusion and future work are explained in Section 3.4.

## 3.2. Proposed Methods

An overview of the proposed methodology is presented in Figure 3.2, followed by a summary of each step in the process. The cGAN is a conditional variation of the GAN, where the generator is instructed to generate a real sample having specific characteristics rather than a generic sample from full distribution (Isola, Zhu, Zhou, & Efros, 2017a). Therefore, we assumed that the cGAN structure is well-suited to outline the breast density area accurately, especially when the training data is limited, and our experimental results support our hypothesis.

As demonstrated in Figure 3.2, the cGAN network comprises two main networks, generator, and discriminator. In our model, the generator Network G takes a mammographic image and tries to generate a mask image of the areas related to dense tissues (i.e., 0 for non-dense pixels and 1 for dense pixels). The generator network then generates data latterly fed into a discriminator network. Finally, the discriminator D learns a loss function to train this mapping by comparing the ground-truth and the predicted output but observing the input image as a condition to improve the network optimization as proposed in (Isola, Zhu, Zhou, & Efros, 2017b).

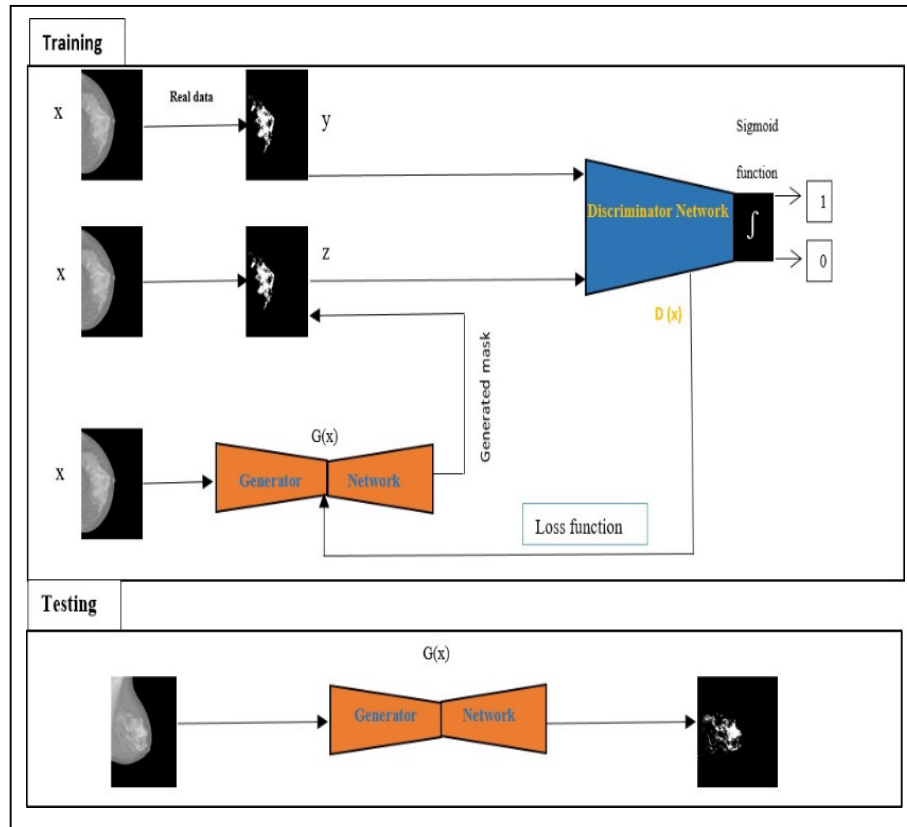


Figure 3. 2 Conditional Generative Adversarial Network framework.

The generator networks follow an encoder-decoder architecture of U-net with skip connection (Ronneberger, Fischer, & Brox, n.d.). The encoder includes down sampling eight convolutional layers. The first layer uses  $7 \times 7$  convolution to generate 64 feature maps, and the final layer generates 512 feature maps with a  $1 \times 1$  size. In contrast, the six middle layers are from the pre-trained ResNet-101 (Ronneberger et al., n.d.). We also used a U-net architecture based on skip connections to improve the segmentation performance. The input of each decoder is concatenated to its corresponding convolutional output of the encoder. In turn, the decoder includes up sampling eight convolutional layers with a reverse ordering layer that is similarly structured to the encoder network.

On the other hand, the discriminator network consists of 5 convolutional layers. The first layer of the discriminator used 64 filters of 3x3 and a stride of 2x2. The final layer of the discriminator produces 512 feature maps with a size of 30x30, followed by Sigmoid as an activation function.

The proposed cGAN model has been trained over a loss function that combines content and adversarial losses. The content loss follows a classical approach in which the predicted dense mask is pixel-wise compared to the corresponding one from the ground truth. In turn, the adversarial loss depends on the real/fake prediction of the discriminator over the ground truth and the predicted foreground mask with observing the input image. The training process of this cGAN can be expressed as an optimization of the following function presented in (3.1), which mathematically describes the training of cGAN.

$$G^* = \mathop{arg\ min}_G \max_D L_{GD}(D, G) + \lambda L_{MSE} \quad (3.1)$$

Where  $\lambda = 10$ , and  $L_{GD}(D, G)$ , the binary cross entropy (BCE) of the adversarial, can be computed as:

$$L_{cGAN}(D, G) = E_{x,y} [\log(D(x, y))] + E_{x,y} [\log(1 - D(x, z))] \quad (3.2)$$

In (2.2), the first term is the entropy of the discriminator D with real data (i.e., the input image is x, and the ground-truth is y, both images are concatenated). The second term is entropy with a fake input data (i.e., the input image is x and the generated image is z, both images are concatenated) passes through the generator, which is then passed through the discriminator to identify the fakeness (i.e., the log probability that the data from generated is fake if it equals to 0). The content loss function computed between z and its corresponding ground-truth y can be defined as:

$$L_{MSE(G)} = \frac{1}{N} \sum_{j=1}^N \|y - z\|_2 \quad (3.3)$$

Where  $N$  is the number of pixels per input image. During training, the discriminator tries to maximize (1), while the task of the generator is precisely the opposite that tries to minimize (1). For our experiments, we have used an Adam (Kingma & Ba, 2014) optimizer with a learning rate of 0.0002, and batch size equals 4, in addition to an optimal number of epochs equals 200.

## 3.3. Experimental results

### 3.3.1. Database

We have used the INbreast dataset (Moreira et al., 2012) to evaluate our results, consisting of 410 mammographic images and presented from MLO and CC views for both sides of the breast (right and left). We manually segmented the 410 images to generate the ground-truth masks assessed by two Hospital Sant Joan de Reus radiologists. This dataset has been divided into three sets, 250 images for training (~60%), 60 images for validation (~15%), and 100 images for testing (~25%).

### 3.3.2. Pre-processing

**Re-sizing:** The mammographic images of the INbreast dataset used in this work have large sizes of 2000x2000 or more. To reduce the computation time, all mammograms were re-sized to a size of  $512 \times 512$  and normalized to scale (0...1).

**Removing pectoral muscle:** The intensity similarity and the overlap between the pectoral muscles and the dense glandular tissues can cause false-positive detection of



dense regions. Pectoral muscle area extraction can help to reduce false positives (Saidin, Mat Sakim, Ngah, & Shuaib, 2013). Hence, we manually removed pectoral muscles of all images in the INbreast dataset using a threshold method proposed (Abdel-Nasser, Moreno, A. Rashwan, & Puig, 2016).

### 3.3.3. Result and discussion

We compared the proposed cGAN to three standard methods of semantic segmentation based on deep learning models fully convolutional network FCN8 (Long, Shelhamer, & Darrell, n.d.), FCN32 (Long et al., n.d.), and Vgg-Segnet (Badrinarayanan, Kendall, & Cipolla, 2017). The three methods mentioned above depend on the generator networks. We computed seven evaluation metrics, recall, specificity, false-positive rate (FPR), false-negative rate (FNR), PBC, precision, and F-score, over the segmented images resulting from the four tested methods. For the best segmentation performance, recall, specificity, precision, and F- score metrics values should be high, however false-negative rate (FNR), false-positive rate (FPR), and PBC should be below.

**Table 3. 1 Recall, Specificity, FPR, FNR, PBC, Precision and F-score with the cGAN-Unet, FCN8, FCN32 and VGG-SegNet methods evaluated on the testing set of the INBreast dataset. The best results are marked in a bold text.**

---

METHODS	RE CALL	SPECIFICITY	PR	NR	BC	PRECISION	F-SCORE
<b>cGAN Unet</b> (Ronneberger et al., n.d.) (Ronneberger et al., n.d.)	<b>0.957</b>	<b>0.999</b>	<b>0.000</b>	<b>0.043</b>	<b>0.022</b>	<b>0.916</b>	<b>0.936</b>
<b>FCN-8</b> (Long et al., n.d.)	0.748	0.997	0.002	0.252	0.506	0.693	0.721
<b>FCN-32</b> (Long et al., n.d.)	0.572	0.997	0.004	0.428	0.795	0.595	0.584
<b>Vgg-Segnet</b> (Badrinarayanan et al., 2017)	0.832	0.996	0.004	0.169	0.558	0.662	0.738

Quantitative measurement results with the testing set (i.e., 100 images) are summarized in Table 3.1. As shown, the GAN-Unet yielded the best results among the four tested methods with an F-score of 94%. In addition, it outperformed all the three methods tested over the six-evaluation metrics. Respectively, the FCN-32 achieved the worst results among the evaluated methods with an F-score of about 50%. This network consists of 32 convolutional layers that need many images to train for proper optimization not available for all cases of medical images datasets. In addition, the SegNet network provided acceptable results with an F-score of 73%.

Qualitative results with the proposed model and the three tested methods are shown in Figure 3.3. Supporting the quantitative results shown in Table 3.1, the segmented images resulting from the cGAN model properly segment the dense regions, including the small regions, preserving the small details and boundaries of the dense tissues. In turn, the other methods, FCN-8, FCN-32, and Vgg-Segnet, smoothed the boundaries of the dense tissues and have not detected the small regions.

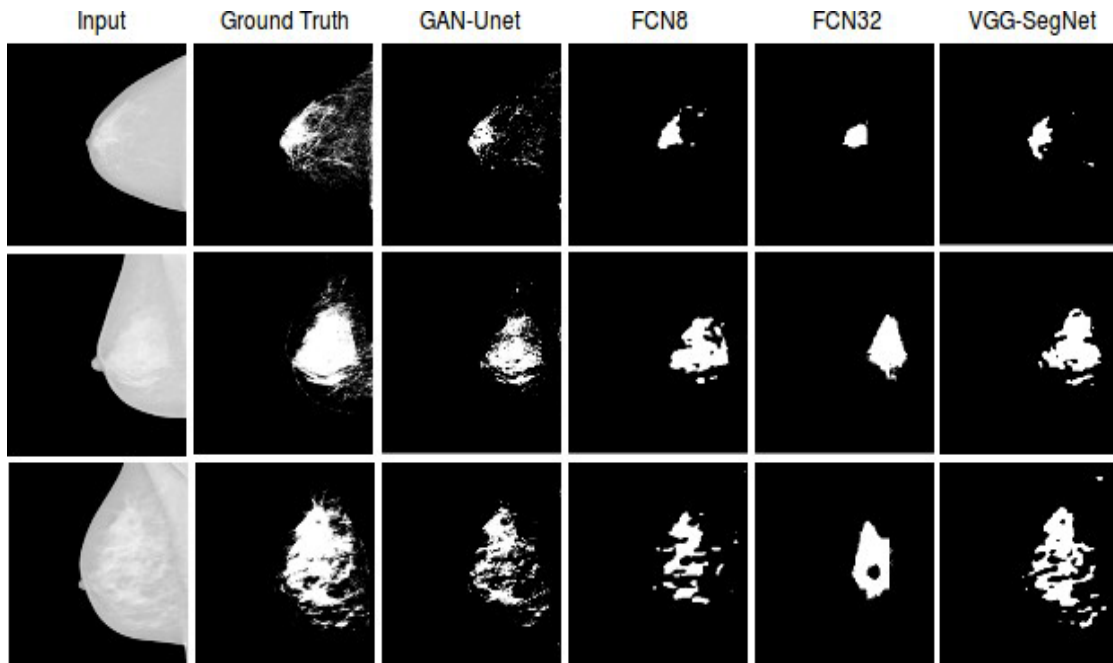


Figure 3. 3 Examples of the breast density segmentation: (col 1) breast mammogram images, (col 2) ground truth mask, (col 3) generated masks with our proposed method cGAN-Unet, (col 4) FCN 8, (col 5) FCN 32 and (col 6) VGG-SegNet.

### 3.4. Conclusion

This chapter proposed to use the conditional Generative Adversarial Network for promising breast density segmentation. The cGAN network consists of two networks. The first network is an encoder-decoder network, called the generator, that maps the input mammogram to a segmented image including the dense areas, and the second network is called discriminator that learns a loss function to enforce the generator to produce an output, which is like the ground-truth. Using the network helps the proposed model to train by a small set of mammographic images. The proposed method can segment the dense regions in mammographic images with overall recall, precision, and F-score around 95%, 92%, and 93%, respectively, outperforming the standard semantic segmentation methods (FCN and SegNet). Future work aims to use the density segmented in mammographic images to generate a more accurate breast density estimation.



# Fully Automated Breast Density Segmentation and Classification using Deep Learning

*Breast density estimation with visual evaluation is still challenging due to low contrast and significant fluctuations in the mammogram's fatty tissue background. Therefore, the primary key to breast density estimation is detecting the dense tissues in the mammographic images. This work intends to develop a fully automated breast tissue segmentation and classification using deep learning. The conditional Generative Adversarial Networks (cGAN) was applied for segmentation; two methods, including thresholding and Convolutional neural network (CNN), were used to perform the breast density classification. For screening mammography, 410 images from the INbreast database were used.*

*The proposed framework can segment the dense regions with the overall accuracy, Dice coefficient, and Jaccard index scores about 98%, 84%, and 72%, respectively. Furthermore, the breast density classification accuracy with thresholding and CNN methods was obtained 92.10% and 99.28%, respectively. Thus, this chapter's findings are promising and show that the proposed technique can produce a clinically applicable computer-aided tool for breast density analysis by digital mammography.*

## 4.1. Introduction

Breast cancer is one of the most common causes of cancer mortality in women worldwide, caused by abnormal cells growing uncontrollably. Those cells may also grow in some places in the human body where they are generally not found. When that happens, the cancer is termed metastatic. Mammography is a standard and most famous radiology tool to detect breast cancer early, possibly before it has outspread. However, investigating these mammographic images is not feasible for every case and highly relies on the radiologist's experience, leading to many false positives. Craniocaudal (CC) and Mediolateral Oblique (MLO) are the most frequent screening mammographic views that provide the best view of the breasts' sidelong, which statistically is the most commonplace for pathological changes. The MLO view is one of the standard views captured from the side of a diagonally compressed breast. In turn, the CC mammographic view is obtained from the superior perspective of a horizontally compressed breast. Various risk factors are associated with breast cancer, of which high breast density is a strong and independent risk factor that can cause developing breast cancer in fibroglandular tissues (Abbas, 2016; Abdel-Nasser, Rashwan, Puig, & Moreno, 2015).

The study of Astley et al. (Astley et al., 2018) shows that the subjective breast density assessment is a more reliable predictor of breast cancer than other automated and semi-automated methods. Besides, whether mammographic density gives rise to more aggressive cancers is debated. Thus, the influence of mammographic density on prognosis should be studied. Mammographic breast density reflects the amount of fibroglandular breast tissue area that appears bright on mammograms, commonly referred to as breast percent density (PD%) (Keller et al., n.d.). In other words, PD% refers to the amount of white or bright area, as seen on a mammogram.

Dense breast tissue is common, non-fatty tissue and is not abnormal, but dense breast tissue can make it harder to identify breast cancer and be associated with an increased risk of breast cancer. It has been reported that women with a high breast density compared to women with a low breast density have a fourfold increased risk of developing the disease (Sprague et al., 2015). Figure 4.1 provides relative risks for developing breast cancer by density category (Sprague et al., 2015).

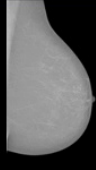
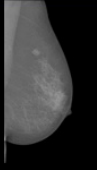
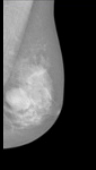
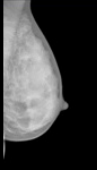
<b>Example</b>				
<b>Category</b>	<b>BI-RADS I</b> Fatty breast	<b>BI-RADS II</b> Average density	<b>BI-RADS III</b> Heterogeneously dense	<b>BI-RADS IV</b> Extremely dense
<b>Range</b>	0-25%	26-50%	51-75%	75-100%
<b>Relative risk</b>	0.5	1	1.5-1.6	1.8-2.0
<b>Prevalence</b>	10%	42%	40%	7%

Figure 4. 1 Prevalence, relative risks of developing breast cancer based on Four classes of the BI-RADS density standard (i.e., fatty, scattered fibroglandular density, heterogeneously dense and extremely dense)

Diverse computational methods have been proposed in the literature for breast density estimation and classification (Boyd, Lockwood, Byng, Trichler, & Yaffe, 1998; Gram, Funkhouser, & Tabár, 1997; Lee & Nishikawa, 2018c; V. A. McCormack, 2006; Wolfe, 1976; Youk, Gweon, Son, & Kim, 2016). For example, to estimate breast density, researchers have proposed methods to segment the dense region of breasts and divide it by the total area of the breast after excluding the pectoral muscle from the MLO

mammograms (Kwok, Chandrasekhar, Attikiouzel, & Rickard, 2006; Tzikopoulos, Mavroforakis, Georgiou, Dimitropoulos, & Theodoridis, 2011). Numerous image segmentation methods have been used for breast density estimation and classification based on handcrafted feature extraction, such as thresholding (Y. Kim, Kim, & Kim, 2010; Nickson et al., 2013), region growing (Nagi, Abdul Kareem, Nagi, & Khaleel Ahmed, 2010; Rouhi et al., 2015), clustering (Keller et al., n.d.), and statistical texture variation (Oliver et al., 2010; Zwiggelaar, 2010).

However, breast density segmentation and classification are still challenging due to the low signal-to-noise ratio and variability of density in appearance and texture (Gandomkar, Suleiman, Demchig, Brennan, & McEntee, 2019a; Matsuyama, Takehara, & Tsai, 2020). Deep Learning (DL), such as Convolutional Neural Networks (CNN), has made several breakthroughs over the past decade, especially in identifying patterns and classifying images. Besides, DL provides several features that other categories of machine learning algorithms do not. Examples of these methods utilized for breast density estimation can be found in the literature (Chan & Helvie, 2019; Lehman et al., 2019). The main contributions of this chapter are as follows.

- I am developing an effective conditional Generative Adversarial Network for segmenting the regions of dense tissues in a mammogram.
- I am utilizing the ratio of the dense segmented regions (i.e., resulting from the cGAN network) to the breast's total area. The computed percentage is used for classifying the mammogram into four different classes of the BI-RADS standard (i.e., fatty, scattered fibroglandular density, heterogeneously dense, and extremely dense).
- I am developing a multi-class CNN architecture for breast density classification using the binary masks obtained from the cGAN.



The rest of this chapter is organized as follows. First, section 2 discusses the related work. Then, section 3 describes the methodology, and Section 4 presents the results. Last, Section 5 concludes the chapter and provides some lines of future work.

## 4.2. Background Study

Various breast density segmentation and classification methods have been proposed in the literature. The techniques that have been employed to perform these approaches include traditional computer vision techniques and deep learning, which will be reviewed in the following sections.

### 4.2.1. Traditional Computer Vision Methods

Traditional Computer-Aided Design (CAD) systems use handcrafted features based on previous experience and radiologists' guidance. Handcrafted feature extraction and breast density classification were initially proposed by Wolfe (*BREAST PATTERNS AS INDEX OF BREAST*, 1976). Wolfe originally described "parenchymal patterns" using verbal descriptions and subjective measures of textural features and classifies breast density into N1—normal fatty breast; P1 and P2—prominent ducts occupying <25% and 25–75% of the breast, respectively; and Dy—dysplastic breast with sheets of dense parenchyma.

In particular, Cumulus software (Byng J W, Boyd N F, Fishell E, Jong R A, 1994) kept growing the ideas of Wolfe (*BREAST PATTERNS AS INDEX OF BREAST*, 1976) by extending the technology and resources to pursue this line of research. Additionally, Cumulus software (Byng J W, Boyd N F, Fishell E, Jong R A, 1994) is an intelligent tool for understanding breast cancer risk, which is a set threshold for segmenting dense tissue,

where regions of breast area are classified into six-category percentages: 0, <10%, 10–25%, 26–50%, 51–75%, and >75%. However, reliance on thresholding might be less accurate, and the significant drawback of threshold-based approaches is that they often lack the sensitivity and specificity needed for accurate segmentation. In this context, five-grade Tábar classification is proposed by Gram et al. (Gram et al., 1997), which classifies breast density into five different categories: I—scalloped contours and Cooper's ligaments, II—evenly scattered terminal ductal lobular units, III—oval-shaped lucent areas, IV—extensive nodular and linear densities, and V—homogeneous structureless fibrosis with convex contours.

Concerning breast density classification methods, well-known breast imaging and reporting data system (BI-RADS) standards have been used (V. A. McCormack, 2006) (Wu, Weinstein, Conant, & Kontos, 2013) to reduce breast imaging evaluation complexity and aid outcome monitoring. BI-RADS classification categorizes breast density into four classes: fatty, scattered fibroglandular, heterogeneously dense, and extremely dense. As shown in Figure 4.1, they could be ranged from almost entirely fatty tissue to extremely dense tissue with very little fat.

Automated techniques include the LIBRA (Laboratory for Individualized Breast Radiodensity Assessment) software (Keller et al., 2012b) based on multi-cluster fuzzy c-means segmentation produced at the University of Pennsylvania, which is publicly accessible. In LIBRA, a total of 86 features are considered, such as global features like Patient's age, breast's thickness X-ray, cluster-merging features like Z-score means, number of unconnected areas, and inter-cluster difference features like compactness and equivalent circular diameter. In addition, LIBRA software (Keller et al., 2012b) also produces area-based analyses of the breast area, dense tissue area, and percentage density from full-field digital mammography (FFDM) images. It is a traditional "handcrafted"

method for breast density estimation, which determined a 0.81 accuracy, but it has many time-consuming and complicated features.

However, the PD% estimated by the algorithm developed in this chapter correlates well with BI-RADS density ratings by radiologists and outperforms LIBRA's algorithm's accuracy. The accuracy achieved by the algorithm developed in this chapter is 0.98 for CC-MLO-averaged, significantly higher than LIBRA's accuracy. The volume-based techniques such as Quantra (Ciatto et al., 2012) and Volpara (Youk et al., 2016) are fully automated software systems to estimate volumetric breast density. Quantra evaluates the thickness of the fibroglandular breast tissue and X-ray attenuation above each pixel in the mammogram images to sum these pixel-wise computations to evaluate the total volume of fibroglandular tissue in the breast. It also evaluates the amount of dense and non-dense tissues at each pixel.

Fully automated methods are currently being developed to obtain a more objective and quantitative breast density evaluation. For instance, the Volpara software (Highnam, Brady, Yaffe, Karssemeijer, & Harvey, 2010; Seo et al., 2013) returns the percentage of dense tissue through a volumetric estimation of the breast. Function only on the raw ("FOR PROCESSING") digital mammogram images, not routinely stored in most medical centers. On the other hand, the semi-automated Cumulus software can display an interactive intensity threshold (Search, Journals, Contact, Iopscience, & Address, n.d.) (Byng J W, Boyd N F, Fishell E, Jong R A, 1994) and is rated one of the best methods for achieving a quantitative segmentation (Boyd et al., 2010). In addition, the Quantra method had a slightly higher, but significantly different, correlation coefficient than the Cumulus method for the volumetric breast density correlation between the right and left breasts ( $r = 0.95$ ,  $p < 0.001$ ).

## 4.2.2. Deep Learning-Based Methods

Recent approaches in machine learning have opened an opportunity to tackle breast density investigating using deep learning methods. Nowadays, deep learning techniques have been used in many studies to automatically extract features from mammograms at multiple levels of abstraction and evidence superior performance. Deep learning networks, such as CNN, can automatically learn features from raw images directly, and they can accurately represent objects at different scales and orientations. CNN is one of the most popular classes of deep neural networks. For example, Kallenberg et al. (Kallenberg et al., 2016) investigated an unsupervised deep learning method based on CNN with four convolutional layers with a max-pooling to learn the characteristics of dense and fatty tissues. Unlabeled imaging data were used to conduct unsupervised feature learning based on CNN to segment the mammogram's breast density regions. The input mammogram is divided into sub-images classified into dense or fatty regions in their approach. The convolutional layers in the unsupervised parts are trained as auto-encoders, and in the supervised part, the (pretrained) weights and bias terms are fine-tuned using Soft-Max regression 5-fold cross-validation (CV). The accuracy of mammographic texture by sparse convolutional autoencoder (MT-CSAE) and the accuracy of mammographic texture density (MT-density) were 0.57 and 0.59, respectively. In another study, Dalmiş et al. (Dalmiş et al., 2017) applied a deep learning-based U-net method for segmenting the breast tissue and achieved the average DSC value of 0.897%. Additionally, the study of Lee and Nishikawa (Lee & Nishikawa, 2018c) proposed a fully convolutional neural network (FCNN) to automatically segment the dense fibroglandular areas on mammographic images. For evaluating their method, 455 full-field digital screening mammograms of 58 cases were used. In addition, they fine-tuned the ImageNet-pretrained VGG16 (Simonyan & Zisserman, 2014) for breast density segmentation and estimation. The Percent Density

(PD) estimation by their approach showed similarities with BI-RADS density assessment by radiologists with 0.81% for CC view, 0.79% for MLO view, and 0.85% accuracy on average. However, automated mammographic breast density estimation using a fully convolutional network (Lee & Nishikawa, 2018c) that applied FCN techniques carried seven convolution layers compared with our methods is more complicated. Aly A. Mohamed. (Mohamed, Luo, Peng, & Jankowitz, 2017) presented a modified Alex Nets for classifying the BI-RADS II and BI-RADS III with accuracy (MLO and CC) = 0.92 and classification the (dense) BI-RADS I and BI-RADS II, (non-dense) (BI-RADS III and BI-RADS IV) with AUC (CC and MLO) = 0.95. In another study, Aly A. Mohamed. (Mohamed, Berg, et al., 2018), a deep learning model using a CNN structure for breast density classification based on the standard categorization of BI-RADS. In addition, two classifiers based on a CNN structure and an improved Alex Net model were proposed in (Mohamed, Luo, Peng, & Jankowitz, 2018) to discriminate the breast density categories. Between categories, BI-RADS II and BI-RADS III, by 6-fold CV and 925 images each, with an accuracy of 94% compared to the local institution's radiological reports. Classification accuracy was reported to increase up to 98% when excluding image data of more inferior quality. Furthermore, the study of Li et al. (Li, Wei, Chan, Helvie, Roubidoux, Lu, Zhou, et al., 2018) presented a technique to separate the breast region into two parts, the "dense region" and the "fatty region." They used deep CNN with three convolution layers, which contained six stages. The first three stages were used as a feature generator, and the second three stages were used as a probability predictor. Their dice similarity coefficient is 0.76%, with a 0.94% correlation coefficient. In the study of Li et al. (Li, Wei, Chan, Helvie, Roubidoux, Lu, Zhou, et al., 2018), the density is classified into only two classes, whereas in the present study, a classifier divided into four classes is proposed. In another study, Lehman et al. (Lehman et al., 2019) have demonstrated deep

convolutional neural network (DCNN) methods based on ImageNet-pretrained ResNet18 for breast density classification. They applied 41,479 images to classify them into two dense and non-dense classes and achieved 0.87% accuracy. Dubrovina et al. (Dubrovina, Kisilev, Ginsburg, Hashoul, & Kimmel, 2018) presented a tissue classification method by supervised CNN framework using a patch-wise approach for CNN training in mammography images. Raw DNN output was recorded as 0.80, and post-processed DNN output was achieved as 0.81. Ciritsis et al. (Ciritsis et al., 2019) applied a DCNN model for breast density classification with 13 convolutional layers, four dense layers in four density classes with two different sets. The first set included 850 MLO, 882 CC, and BI-RADS accuracy was between 71% and 71.7%, dense-vs.-non. dense: 88.6–89.9%. The second set included 100 MLO and 100 CC; the accuracy of BIRADS was between 87.4% and 92.2%, dense-vs.-non. dense: 96–99%. Gandomkar et al. (Gandomkar et al., 2019a) investigated Inception-V3 network architecture to process the mammograms and pretrained the system based on ImageNet. Their network achieved an accuracy of 92.0% in high against low-risk classification.

Many breast density classification methods in mammograms have been presented in the literature, but only a few studies have achieved accuracy above 90%, which is more complicated than the method presented in this chapter. In this chapter, a novel method of representing breast tissue is presented. We are modeling dense tissue distribution to fatty tissue and how this can be utilized to provide the density segmentation classification based on BI-RADS and density percentage. It should be noted that all mammograms were classified according to the BI-RADS density classification system by expert breast radiologists. The novel and advantageous features of the method proposed in this chapter are as follows.

- The first adaptation of cGAN in fully automated breast density segmentation in mammograms is developed,
- The breast density percentage classification by the developed multi-class CNN architecture correlated well with BI-RADS density ratings (BI-RADS I, BI-RADS II, BI-RADS III, and BI-RADS IV) using the binary mask segmented in the previous stage (cGAN output) by radiologists,
- A strong correspondence between the output of our automated algorithm and radiologist has presented breast density measures can be obtained, and
- The proposed approach results in remarkably faster calculation while improving the classification efficiency compared to other methods in the literature

### 4.3. Methodology

This section gives detailed information about the methodology used in this chapter. The full research methodology is shown in Figure 4.2, and it is divided into two stages: stage one includes breast mammogram segmentation into background, pectoral muscle, and breast tissue region. The second stage corresponds to breast density classification based on BI-RADS.

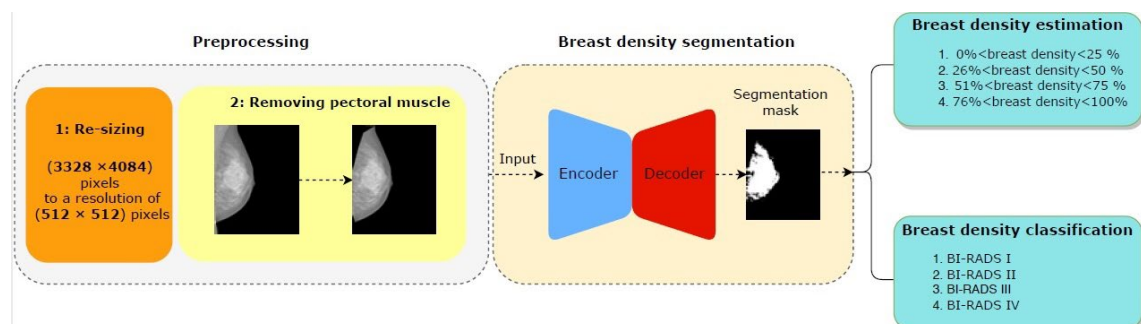


Figure 4. 2 An overview of the proposed framework

In the first stage, mammograms are preprocessed for removing pectoral muscles. Later, the cGAN input is rescaled by resizing the mammograms to  $512 \times 512$  pixels, including different breast densities. The processed mammograms are then fed to the proposed cGAN to get a binary mask containing the dense tissue. In the second step, two methods are used: (1) computing the percentage of breast density by dividing the area's dense tissues into the breast's total area, and (2) the output binary mask is down-sampled into  $128 \times 128$  pixels, which is used to classify breast density in 2 different ways; first, the output of the binary mask is fed to a multi-class CNN to classify the breast density into four classes based on BI-RADS, and second, the breast density percentage by the traditional method based on thresholding rules is estimated.

Several techniques are applied in this chapter to prepare the dataset before feeding the mammograms into the proposed model; the stepwise details are explained in the following subsection.

### **4.3.1. Preprocessing**

Some preprocessing operations in the first step are used, such as removing the pectoral muscles and resizing the mammographic images. However, the high similarity in intensity and the overlap between the pectoral muscle and the glandular tissue can cause false-positive detection of dense tissue areas in mammographic images. Therefore, extraction of the pectoral muscle area is applied, reducing the false positives (Seo et al., 2013), as the identification and removal of pectoral muscles play an essential preprocessing step in the CAD system (Kallenberg et al., 2016). An automatic method described in our preliminary work (Abdel-nasser, Moreno, & Puig, 2016) has been utilized to remove the pectoral muscles. This method involves three main steps. First, the breast region and the pectoral muscle are segmented from the background, and then secondly,



the mammogram orientation is determined. Finally, a region growing segmentation is used for removing the pectoral muscle from the image. The example of removing the pectoral muscles is shown in Figure 4.3. In addition, to reduce the computation time, all mammograms were resized from  $(2560 \times 3328)$  or  $(3328 \times 4084)$  pixels to a resolution of  $(512 \times 512)$  pixels (i.e., the resolution yields the best accuracy for the segmentation stage).

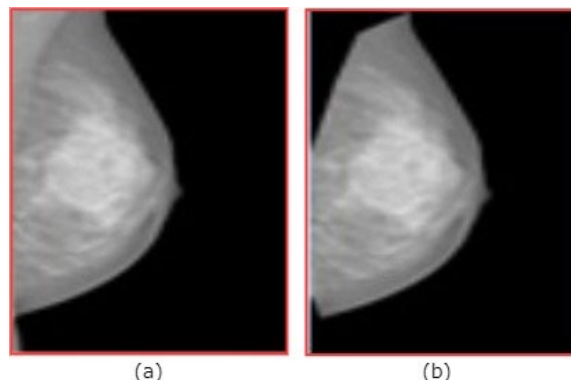


Figure 4. 3 (a) Original images. (b) After removing the pectoral muscle.

### 4.3.2. Breast Density Segmentation

The proposed framework of breast density segmentation is presented in Figure 4.4, followed by a summary of each step in the process. For dense tissue segmentation, the conditional cGAN is used, which is proposed by (Isola et al., 2017a). cGAN is a conditional variation of the GAN, where the generator is instructed to generate a real sample having specific characteristics rather than a generic sample from full distribution (Byng J W, Boyd N F, Fishell E, Jong R A, 1994). It has been assumed that the cGAN

structure is well suited to accurately outline the breast density area, especially when the training data is limited, and our experimental results support our hypothesis. As demonstrated in Figure 4.4, the cGAN network comprises two main networks: generator and discriminator.

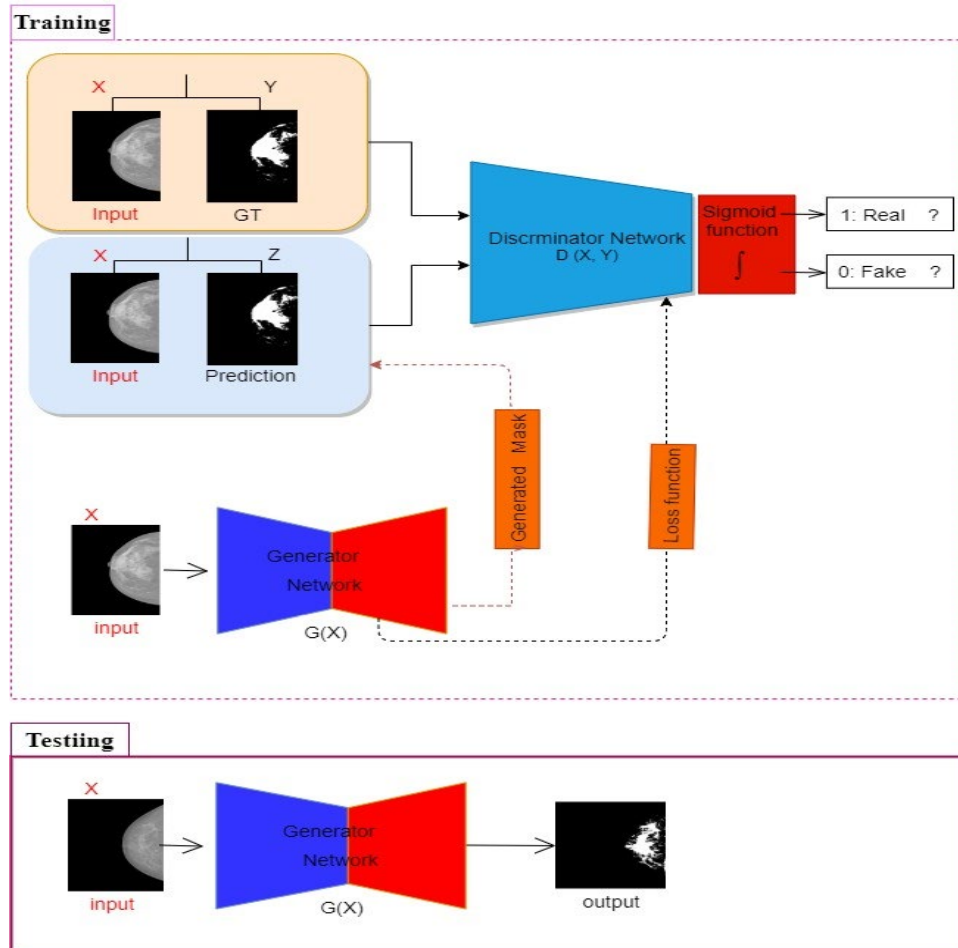


Figure 4. 4 cGAN framework for segmentation.

The Generator network  $G$  comprises two parts: encoder and decoder layers. The encoder layers help extract the input images' features (e.g., texture, edge, shape, and intensity). On the other hand, decoder layers generate a binary mask according to these extracted features. The Discriminator network  $D$  works as a classifier to discern the generated binary mask and its corresponding ground truth. This adversarial network consistently enforces the generator network by its working methodology during the training process. In the model presented in this chapter, the  $G$  network takes a

mammographic image and tries to generate a mask image of the areas related to dense tissues (i.e., 0 for non-dense pixels including the background pixels, and 1 for dense tissues pixels). The generator network then generates data latterly fed into a discriminator network. Finally, the discriminator  $D$  learns a loss function to train this mapping by comparing the ground-truth and the predicted output but observing the input image as a condition to improve the network optimization as proposed (Isola et al., 2017a). The  $G$  network follows an encoder-decoder architecture of U-net with skip connection (Ronneberger et al., n.d.). The encoder includes down sampling eight convolutional layers. The first layer uses  $7 \times 7$  convolution to generate 64 feature maps, and the final layer generates 512 feature maps with a  $1 \times 1$  size. At the same time, the six middle layers are from the pretrained ResNet-101 (Ronneberger et al., n.d.). In turn, the decoder includes up sampling eight convolutional layers with reverse ordering layers that are similarly structured to the encoder network. A U-net architecture based on skip connections in which each decoder's input is concatenated to its corresponding convolutional output of the encoder is also used in this chapter to improve the segmentation performance.

On the other hand, the discriminator network consists of 5 convolutional layers. The first layer of the discriminator used 64 filters of  $3 \times 3$  and a stride of  $2 \times 2$ . The final layer of the discriminator produces 512 feature maps with a size of  $62 \times 62$ , followed by Sigmoid as an activation function. The proposed cGAN model has been trained over a loss function that combines content and adversarial losses. The content loss  $L_c$  follows a classical approach in which the predicted dense mask is pixel-wise compared to the corresponding one from the ground truth. Three loss functions are tested for this loss: Mean Square Error (MSE), Dice, and structural similarity index (SSIM). The adversarial loss depends on the real/fake prediction of the discriminator over the ground-truth and the predicted foreground mask with observing the input image. For example, assume the input

mammography image is  $x$ ,  $y$  the ground truth mask,  $z$  a random variable,  $\lambda$  an empirical weighting factor,  $G(x, z)$  and  $D(x, G(x, z))$  the outputs of  $G$  and  $D$ , respectively. Thus, the training process of this cGAN can be expressed as an optimization of the following objective function presented in (4.1), which mathematically describes the training of cGAN.

$$G^* = \arg \min_G \max_D L_{cGAN}(D, G) + \lambda L_c(G) \quad (4.1)$$

where  $\lambda=10$ , and  $L_{cGAN}(D, G)$ , the binary cross-entropy (BCE) of the adversarial, can be computed as shown in (4.2):

$$L_{cGAN}(D, G) = E_{x,y} (\log(D(x, y))) + E_{x,z} \log(1 - D(x, z)) \quad (4.2)$$

In (4.2), the first term is the entropy of the discriminator  $D$  with real data (i.e., the input image is  $x$ , and the ground-truth is  $y$ , both images are concatenated). The second term is entropy with a fake input data (i.e., the input image is  $x$  and the generated image is  $z$ , both images are concatenated) passes through the generator, which is then passed through the discriminator to identify the fakeness (i.e., the log probability that the data from generated is fake if it equals to 0), and the content loss function computed between  $z$  and its corresponding ground-truth  $y$ , can be defined using (4.3):

$$L_c(G) = f(y, G(x, z)) \quad (4.3)$$

Where  $f$  is MSE, Dice, or SSIM loss functions. The MSE loss function can be computed using (4.4):

$$L_{MSE}(y, G(x, z)) = \frac{1}{N} \sum_{j=1}^N \|y - G(x, z)\|_2 \quad (4.4)$$

where  $N$  is the number of the pixels per input image and  $L_{Dice}(y, G(x, z))$  is the dice loss of the predicted mask concerning ground truth, which is defined using (4.5):

$$L_{Dice}(y, z) = 1 - \frac{2|y \times G(x, z)|}{|y| + |G(x, z)|} \quad (4.5)$$

The SSIM (Z. Wang, Bovik, Sheikh, & Simoncelli, 2004) considers contrast, luminance, and structure to determine the similarity between two images. SSIM can be calculated using (4.6) and (4.7).

$$SSIM(y, z) = 1 - \frac{(2\mu_y\mu_z) + c_1(2\sigma_{yz} + c_2)}{(\mu_y^2 + \mu_z^2 + c_1) + (\sigma_y^2 + \sigma_z^2 + c_2)} \quad (4.6)$$

$$\sigma_{yz} = \frac{1}{T-1} \sum_{i=0}^T (y_i - \mu_y)(z_i - \mu_z) \quad (4.7)$$

where  $y$  and  $z$  are the ground-truth and generated images, respectively;  $\mu_y$  is the average of  $y$ , and  $\mu_z$  is the average of  $z$  (are local measures of the mean of the ground-truth and generated images).  $\sigma_y$  is the standard deviation of  $y$ ,  $\sigma_z$  is the standard deviation of  $z$ .  $\sigma_{yz}$  is the local measures of the correlation between two images,  $\sigma_y^2$  and  $\sigma_z^2$  are the local measures of the variance of the two images; and  $C_1=(k_1L)^2$ ,  $C_2=(k_2L)^2$  are some predefined constants that are two variables to stabilize the division with small denominators,  $L$  is the dynamic range of the pixel values (typically this is 255),  $k_1 = 0.01$  and  $k_2 = 0.03$  by default.  $T$  is the total number of pixels in each image. The optimization process of  $G$  tries to minimize both expected values, i.e., the  $D$  values should approach 1.0 (correct tumor segmentation), and the content loss  $L_c$  should approach 0.0 (generated masks equal to ground truth). Both terms of generator loss enforce the proper optimization of  $G$ : the dice loss term fosters a rough prediction of the mask shape (central tumor area), while the adversarial term fosters an accurate prediction of the mask outline (tumor borders). Neglecting either of the two terms may lead to very poor segmentation results or

slow learning speed. During the training process, the discriminator tries to maximize the function presented in (4.1), while the task of the generator is precisely the opposite that tries to minimize the function presented in (4.1).

For our experiments, an Adam optimizer (Byng J W, Boyd N F, Fishell E, Jong R A, 1994) with a learning rate of 0.0002 and batch size equal to 4 and an optimal number epochs equals 200 has been used. Two main training and testing procedures can be distinguished. First, during training in a supervised mode, the classifier learns to distinguish between fatty and dense pixels from manually annotated images, whereas in testing, the classifier assigns a fatty or dense tissue label to each pixel of the input image (*BREAST PATTERNS AS INDEX OF BREAST*, 1976). Figure 4.4 shows the framework of the proposed method for breast density segmentation. Moreover, a pretrained ResNet-101 was used as a base feature extractor, illustrated in Figure 4.5.

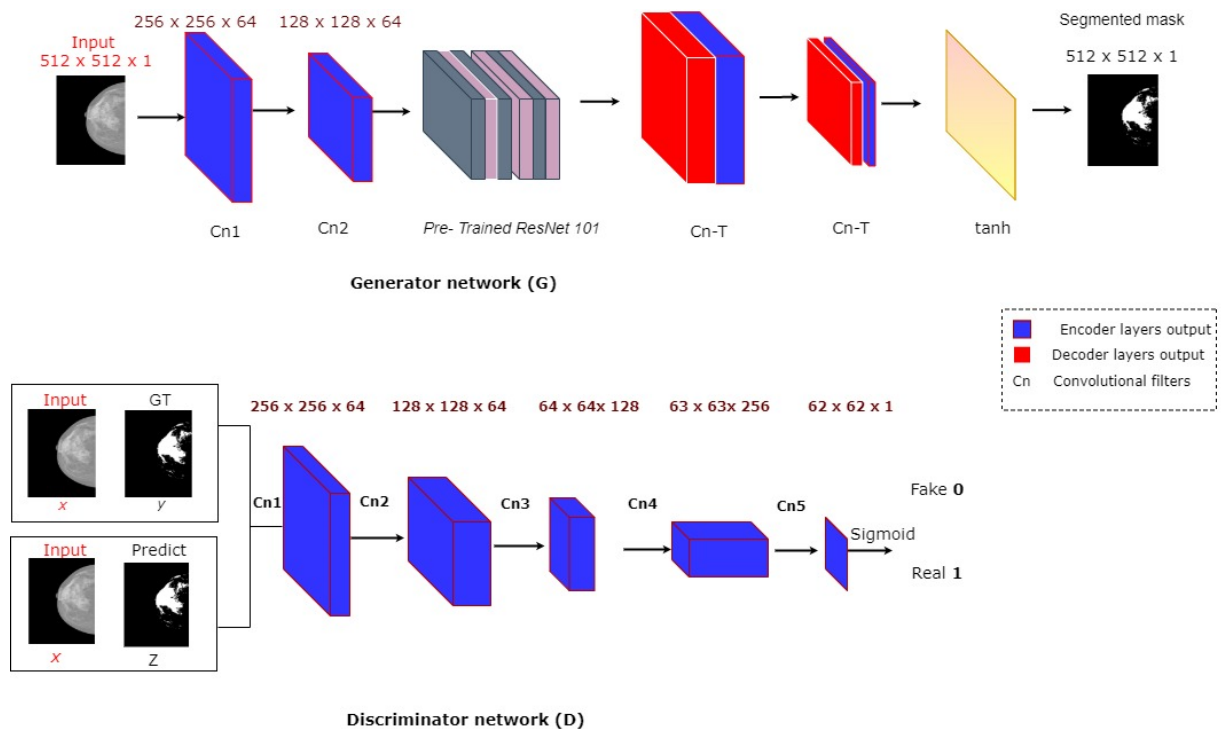


Figure 4. 5 cGAN framework for breast density segmentation.

### 4.3.3. Breast Density Classification

Two techniques, including a traditional method and a CNN-based method, have been used; each method has four output classes described in the following sections.

#### 4.3.3.1. Breast Density Percentage Estimation Based on Traditional Method

Percent density obtained from mammographic images refers to the ratio of the area of dense tissue present in a mammogram to the total area of the breast. For the traditional method, we perform five stages:

- First, we resize the generated mask images to the same resolution as the input mammography.
- We express the breast region area by the number of non-zero pixels in the mammogram images.
- We then count the non-zero pixels in the generated mask for expressing the area-dense tissues.
- Computing the ratio between the area of dense tissues and the area of the breast region to estimate the breast density in the input image.
- Finally, by the thresholding rules defined in the BI-RADS density standard, we classify the breast density to 4 categories: ( $0\% < \text{BDE} < 25\%$ ), ( $26\% < \text{BDE} < 50\%$ ), ( $51\% < \text{BDE} < 75\%$ ), ( $76\% < \text{BDE} < 100\%$ ).

#### 4.3.3.2. Breast Density Classification Based on a CNN

Figure 4.5 shows the CNN architecture for breast density classification (four classes correspond to the BI-RADS categories: fatty, scattered fibroglandular density, heterogeneously dense, and extremely dense) using the binary masks obtained from the cGAN.

Most methods that are attempted to categorize breast density have computational complexity. As shown in Figure 4.6, the proposed CNN technique consists of three convolutions layers with kernel sizes  $9 \times 9$ ,  $5 \times 5$ , and  $4 \times 4$ , respectively, and two fully connected (FC) layers. The first two convolutions' layers are followed by  $4 \times 4$  max-pooling with stride  $4 \times 4$ . The last convolution layer's output is flattened and fed into the first FC layer with 128 neurons. These four layers use ReLU as an activation function. A dropout of 0.5 is used to reduce overfitting in the first FC layer. Finally, the last FC layer with four neurons applies the soft-max function to generate the input binary mask's final membership degree to each class. A weighted categorical cross-entropy loss is used to avoid the problem of an unbalanced dataset. The class weight is one minus the ratio of samples per class to the total number of samples.

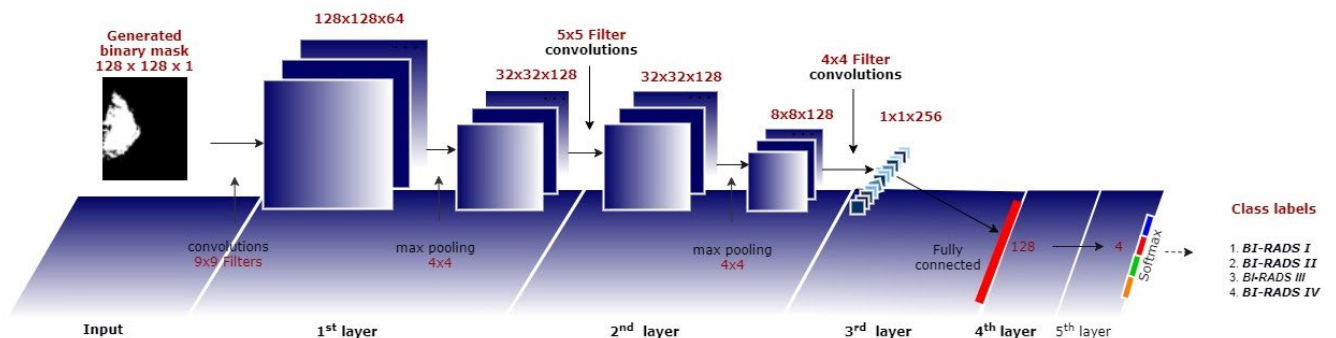


Figure 4.6 Proposed CNN architecture used for breast density classification (second technique).

The RMSProp is applied for optimizing the model with a learning rate of 0.001, a momentum of 0.9, and a batch size of 16. The network is trained from scratch, and the weights of 5 layers are randomly initialized. The best architecture, number of layers, filters per layer, and neurons in FC layers during training were found experimentally.



## 4.4. Experimental Results

### 4.4.1. Dataset

INbreast dataset<sup>2</sup>: The initial results presented in this section are based on the INbreast dataset (Moreira et al., 2012). It is a publicly available 2-dimensional (2D) database and includes MLO and CC mammographic images of 115 patients (410 mammograms). Thus, every patient has four mammographic images: MLO-right, MLO-left, CC-right, and CC-left.

It has a ground truth for mass location, mass type, and breast density classification label. Breast density classification in INbreast was prepared based on breast imaging reporting and data method (BI-RADS) standard (BI-RADS I, BI-RADS II, BI-RADS III, and BI-RADS IV). The image size of a mammogram is  $3328 \times 4084$  or  $2560 \times 3328$  pixels. Note that the INbreast dataset does not have the ground truth binary masks for the breast density segmentation. Thus, we have annotated the images with the cooperation of radiologists' experts in breast cancer. The INbreast dataset has 115 patients divided into 82 patients (80% of the total images) as a training set and 33 patients (20% of the total images) as a test set. To train the proposed CNN network, we applied the "Holdout cross-validation method" on the images of the 82 patients to divide it into 80% for training and the rest for validation. The "Holdout cross-validation method" ensures that the images are randomly divided into training and validation sets without any intersection between them to guarantee a fair evaluation.

<sup>2</sup> [http://medicalresearch.inescporto.pt/breastcancer/index.php/Get\\_INbreast\\_Database/](http://medicalresearch.inescporto.pt/breastcancer/index.php/Get_INbreast_Database/)

In the holdout cross-validation, (33%) of data belong to BI-RADS I with 108 images of 27 patients, (35%) of data belong to BI-RADS V with 116 images of 29 patients, (25%) of data belong to BI-RADS III with 80 images of 20 patients, and only (7%) of data belong to BI-RADS IV with 22 images of 6 patients. Table 4.1 shows the INbreast dataset distribution for training and test sets.

Table 4. 1 Distribution of densities across imbalanced INbreast dataset (before augmentation).

	I	II	III	IV
<b>Total dataset</b>	33%, 136 images	35%, 146 images	25%, 100 images	7%, 28 images
<b>(410 images)</b>	From 38 patients	From 41 patients	From 25 patients	From 8 patients
<b>Training and validation subset</b>	33%, 108 images	35%, 116 images	25%, 80 images	7%, 22 images
<b>(328 images)</b>	From 27 patients	From 29 patients	From 20 patients	From 6 patients
<b>Test subset</b>	33%, 27 images	35%, 29 images	25%, 20 images	7%, 6 images
<b>(82 images)</b>	From 11 patients	From 12 patients	From 5 patients	From 2 patients

One of this chapter’s principal goals is to create a robust model that generalizes well to new data and uses images of patients not initially included in the training stage. For testing, 33 patients have been used as a control/test set for evaluating the performance of the trained deep models for dense tissue segmentation and breast density classification. Thus, our control/test set can serve as a proxy for new data. Table 4.1 breaks down breast density variety from the INbreast dataset. The distribution of breast densities variety across the four classes is shown in Table 4.1. As shown, the dataset is highly imbalanced, with the lowest percentage (7%) of data belonging to BI-RADS IV and the highest percentage of data (35%) belonging to BI-RADS II.

Indeed, the INbreast dataset used for training the segmentation model does not have the ground truth (i.e., binary masks) for the dense tissue segmentation. However, it

annotates the images with the corresponding class of BI-RADS. The lack of ground truth for breast density assessment is a limitation. Therefore, the dense regions in the INbreast dataset images are segmented by two radiologists from Hospital Sant-Joan de Reus (Reus Sant-Joan Hospital located in Tarragona province, Spain). Pixel-wise logical-AND was applied on the binary masks generated by the two radiologists to generate the ground truth, meaning the two radiologists have to agree about the dense tissues in the same mammogram. Assume we have a binary mask for each mammogram generated by each radiologist. If the same pixel has a value of 1 in both binary masks, it will be 1 in the final ground-truth image.

#### 4.4.2. Implementation Details

The proposed method was applied using Python v.3.5 with PyTorch<sup>3</sup> library, running on a 64-bit Ubuntu operating system, a 3.4 GHz Intel Core-i7 CPU with 16 GB of RAM, and NVIDIA GTX 1070 GPU with 8 GB of video RAM.

#### Evaluation Metrics of Breast Density

The terms and formulas involved in evaluating the results of breast density classification are described in (3.8) to (3.11):

$$Accuracy = \frac{\text{Correct predocions}}{\text{Total predictions}} = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (4.8)$$

$$Precision = \frac{\text{True positive}}{\text{predicted positive}} = \frac{TP}{(TP + FP)} \quad (4.9)$$

$$\text{Sensitivity} = \text{True – PositiveRate} = \text{Recall} = \frac{\text{True positive}}{\text{Actual positive}} = \frac{TP}{TP + FN} \quad (4.10)$$

$$\text{Specificity} = \text{True – Negative Rate} = \frac{TN}{TN + FP} \quad (4.11)$$

<sup>3</sup> <https://pytorch.org>

This part needs to compute TP, FN, FP, and TN for a multi-class problem with one score for each class and count any other class as a negative. For example, in our case, we have four classes (1, 2, 3, 4); thus, TP, FN, FP, and TN for C1 can be calculated as True positive (TP) instances are gold standard class 1 predicted as class 1 False-negative (FN) instances are gold standard class 1 predicted as class 2, 3, or 4 False-positive (FP) instances are gold standard class 2, 3, or 4 predicted as class 1 True negative (TN) instances are gold standard class 2,3 or 4 predicted as class 2, 3, or 4 (here errors do not matter as long as class 1 is not involved) Similarly, for the three other classes C2, C3, and C4, we can compute TP, FN, FP, and TN.

### **4.4.3. Breast Density Segmentation**

It is crucial to accurately estimate the breast density to achieve proper dense tissue segmentation. In the first experiment, three variations of the cGAN-Unet network with the different content loss function  $L_c$  are evaluated: MSE, Dice, and SSIM. For the quantitative analysis, the dense region's segmentation quality is measured using three evaluation metrics: accuracy, Dice coefficient or F1 score (DSC), and Jaccard index (JI). Quantitative results are shown in Table 4.2. Note that we separately computed the metric value of the images of class 1 (C1), class 2 (C2), class 3 (C3), and class 4 (C4), while "all" refers to the metric value of all images of the testing set. Table 4.2 summarizes all the methods tested over the three-evaluation metrics: Accuracy, Jaccard index, and Dice over the segmented images.

Table 4. 2 Accuracy, DSC, and JI with the cGAN-Unet, cGAN-Unet-SSIM-loss, and cGAN-Unet-dice-loss evaluated on the testing set of the INBreast dataset for breast density segmentation (C1 = Class1, C2 = Class2, C3 = Class3, C4 = Class4).

Model	Accuracy					DSC					J I				
	C1	C2	C3	C4	all	C1	C2	C3	C4	all	C1	C2	C3	C4	all
<b>cGAN-Unet-Dice. loss</b>	0.98	0.99	0.99	0.99	0.98	0.66	0.90	0.95	0.95	0.88	0.50	0.82	0.91	0.91	0.78
<b>cGAN-Unet-SSIM. loss</b>	0.94	0.98	0.98	0.96	0.96	0.54	0.86	0.91	0.85	0.79	0.37	0.75	0.83	0.74	0.65
<b>cGAN-Unet-MSE. loss</b>	0.68	0.84	0.93	0.95	0.80	0.53	0.85	0.94	0.97	0.81	0.36	0.74	0.90	0.94	0.67

As shown in Table 4.2, the cGAN-Unet with Dice provides the best accurate dense regions segmentation, among the other proposed models, with an accuracy of 98%. cGAN-Unet with Dice yields a significant improvement of 7% and 11% with DSC and JI, respectively, compared to the cGAN-Unet with MSE, which reflects the highest similarity between the ground truth and the predicted segmentation. The segmentation performance of cGAN networks for each class of the INbreast dataset, using Dice, gives the best results for the four classes inaccuracy. In turn, cGAN-Unet with an SSIM loss is the second-best model yielding an accuracy of 96%, DSC of 79%, and JI of 65%). The lowest overall dense tissue segmentation performance with an accuracy of 80% has been obtained by cGAN-Unet using the MSE content loss function, in which the values of DSC and Jaccard scores achieved are 80% and 67%, respectively. The results with the three variations indicate that DSC and SSIM as loss functions help the adversarial network train the generative network better than MSE. The poor performance of the MSE loss function is because it is prone to outliers. After all, it uses the Mean in computing each error value. In turn, the DSC and SSIM loss increases the similarity between the dense segmented regions and the ground truth. Figure 4.7 supports the quantitative results of Table 4.2, as the segmented images resulted in the cGAN-Unet with a Dice loss accurately segmented the dense regions, including the small regions and preserving the small details and boundaries of the dense tissues. To assess the proposed model (i.e., cGAN-Unet with a Dice loss providing the best

results in Table 4.2) and to show its effectiveness, it was compared against state-of-the-art segmentation models that are commonly used for semantic segmentation based on deep learning models such as FCN8 (Long et al., n.d.), FCN32 (Long et al., n.d.), and Vgg-Segnet (Badrinarayanan et al., 2017).

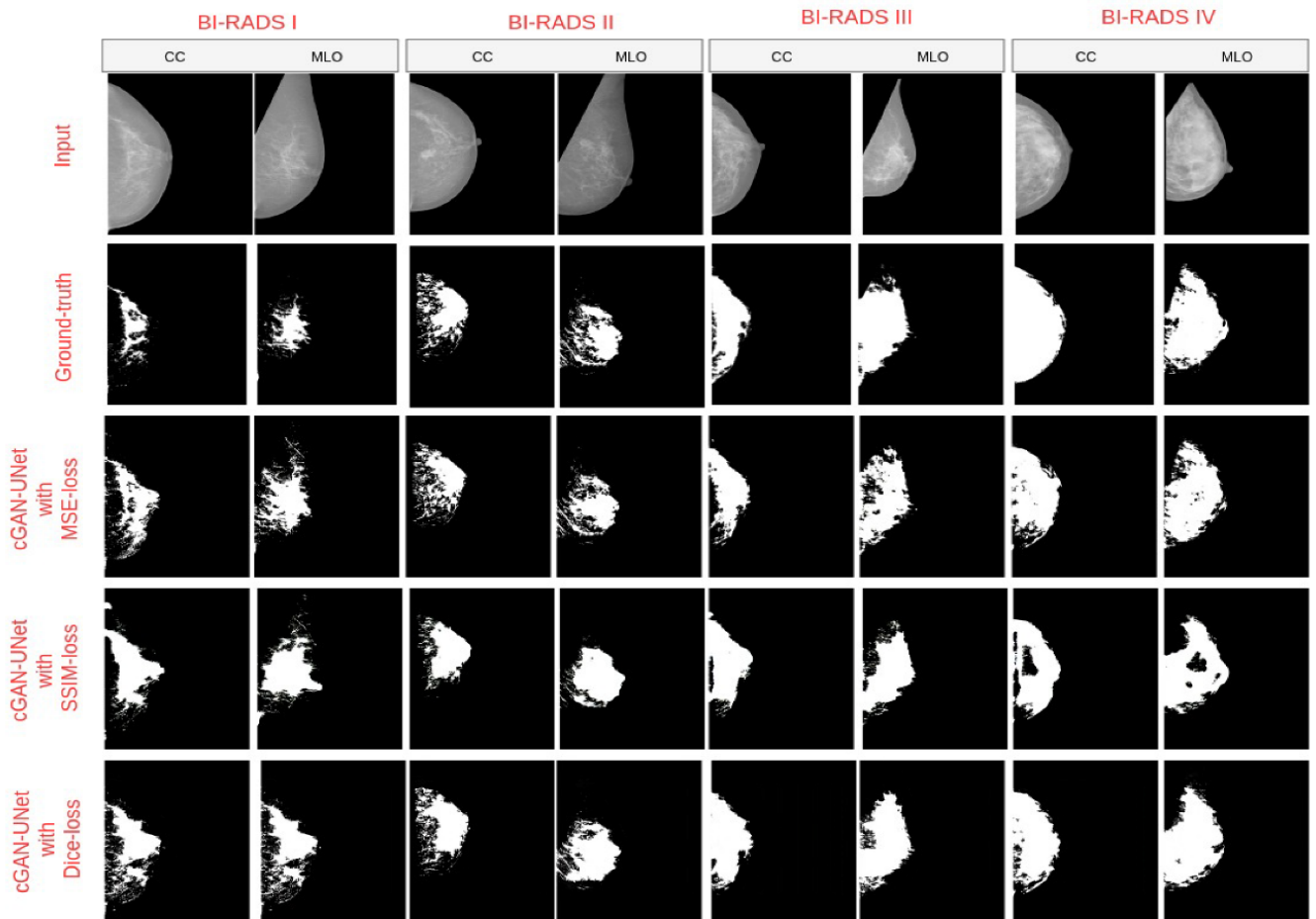


Figure 4. 7 Breast density segmentation result of three models with the INbreast dataset. (Row 1) original images. (Row 2) Ground truth. (Row 3) Result of cGAN-UNet with dice-loss. (Row 4) Result of cGAN-UNet with myssim loss. (Row 5) Result of cGAN-UNet. (Col 1) BI-RADS I from two different views (CC and MLO). (Col 2) BI-RADS II. (Col 3) BI-RADS III. (Col 4) BI-RADS IV.

The results of this assessment are shown in Table 4.3. All models are trained and tested on the INbreast dataset. It is noteworthy that these results show that the proposed model developed in this chapter outperformed the other models in terms of Sensitivity, Specificity, Precision, and DSC score. On the other hand, the FCN-32 achieved the worst results among the evaluated methods with a DSC score of 58%. This network consists of

32 convolutional layers that need many images to be appropriately trained; however, most medical datasets lack enough images as the main difficulty. Therefore, the FCN-8 model with eight layers improved 14% in the DSC score compared to FCN-32. Besides, the Vgg-Segnet network provided acceptable results with a DSC of 73%. In turn, cGAN-Unet with the Dice loss improves 15% in the DSC score compared to the Vgg-Segnet.

Furthermore, it yields an improvement of 12% of Precision better than the best second method, FCN-8. Regarding Sensitivity, cGAN-Unet achieved the best result among the four deep models with an improvement of 12% compared to Vgg-Segnet, the second-best model. For Specificity, the four methods provide very high Specificity; however, our model based on cGAN yields the lowest Specificity of 98.5%, among them only 1.2% lower than FCN-8 and FCN-32 models.

Table 4. 3 Summary of Sensitivity, Specificity, FPR, FNR, Precision, and DSC with the cGAN-Unet-MSE-loss, FCN8, FCN32, and VGG-SegNet methods evaluated on the testing set of the INBreast dataset for breast density segmentation

Methods	Sensitivity	Specificity	Precision	DSC
<b>cGAN-Unet with Dice-loss</b>	<b>0.957</b>	<b>0.985</b>	<b>0.81</b>	<b>0.88</b>
FCN-8 (Shelhamer, Long, & Darrell, 2017)	0.748	0.997	0.69	0.72
FCN-32 (Shelhamer et al., 2017)	0.5724	0.997	0.59	0.58
Vgg-Segnet (Badrinarayanan et al., 2017)	0.832	0.996	0.66	0.73

In conclusion, cGAN-Unet, with its variations proposed in this chapter, can learn the statistical invariant features (texture, color, etc.) of an input image and then generate nearly segmented images, which look like the ground-truth image. However, this chapter's segmentation model contains about 13,607,043 parameters for tuning the generator part in the cGAN network. The method developed in this chapter is fast in both training, i.e., around 30 s per epoch (326 images), and predicting, around seven images per second. Surprisingly, that is 10–15 times faster than the FCN-32 mode and 7–8 times faster than the FCN-8 model.

#### 4.4.4. Breast Density Classification

This chapter proposed two different classification techniques with four output classes. In the first technique, the traditional breast density method is classified into four categories, and thresholding rules estimate breast density percentage. Whereas in the second method, a CNN model is applied on the generated binary masks. A confusion matrix of the traditional method is provided in Table 4.4. For the traditional method, the overall accuracy of breast density percentage classification based on thresholding rules is 80% for class-1 (BI-RADS I); the traditional method correctly classified 77% of the images. The traditional method of breast density gives the lowest accuracy with class 2. This happens as there is a high correlation between class-1 and class-2. For class-3 (BI-RADS III) and class-4 (BI-RADS IV), we achieved classification rates of 90% and 84%, respectively.

**Table 4.4 Confusion matrix of breast density estimated based on thresholding rules.**

Class	Predicted label			
	I	II	III	IV
Ground truth				
I	0.77	0.20	0.03	0
II	0	0.76	0.24	0
III	0	0	0.90	0.10
IV	0	0	0.16	0.84

For the proposed breast density classification based on a trained CNN network, two experiments have been applied: one for the classification of the imbalance dataset (The training data without augmentation) and the other one for a balanced dataset (the training data after augmentation). The network has been tested with two different image sizes, 64×64 and 128×128. Note that the input to the CNN-based method is the binary image generated by the segmentation model. As explained in the above subsection of evaluating



breast density segmentation, the cGAN-Unet with a dice loss yields the best segmentation results. The results of this network were used to classify the breast density in mammography. Due to the imbalanced data set, deep learning classification performance may decrease. Thus, to overcome the imbalanced number of training images, we have done data augmentation by applying “Illumination change,” “scaling,” and “flipping,” which yields 798 images for each class, as shown in Table 4.5. We applied the augmentation techniques on the training set only with 326 images of 82 patients. In turn, the same test set of 84 images of 33 patients is used for evaluating the two trained CNN networks.

Table 4. 9 Distribution of densities across balanced INbreast dataset (after augmentation).

	I	II	III	IV
Total dataset (3192 images)	25% (798 images)	25% (798 images)	25% (798 images)	25% (798 images)
Training subset (2552 images)	25% (638 images)	25% (638 images)	25% (638 images)	25% (638 images)
Validation subset (640 images)	25% (160 images)	25% (160 images)	25% (160 images)	25% (160 images)
Test subset (82 images)	27 images of 11 patients	29 images of 12 patients	20 images of 5 patients	6 images of 2 patients

The evaluation of CNN classification accuracy, precision, sensitivity, and specificity has been applied on imbalanced and balanced datasets with two different input images ( $64 \times 64$  and  $128 \times 128$ ) detailed in Table 4.6.

Table 4.6 Accuracy, Precision, Sensitivity, and Specificity of the CNN-based classification method on imbalanced and balanced datasets with two different sizes of input images:  $64 \times 64$  and  $128 \times 128$ .

	Size of Input Images	Accuracy	Precision	Sensitivity	Specific
Imbalanced dataset (410 images)	$128 \times 128$	90.29	90.29	90.29	96.76
	$64 \times 64$	94.95	94.17	94.17	98.05
Balanced dataset (3192 images)	$128 \times 128$	98.75	97.5	97.5	99.16
	$64 \times 64$	98.62	97.85	97.85	99.28

As shown in Table 4.6, the CNN-based classification method's performance is higher when using a balanced dataset than the imbalanced data set in terms of the four evaluation measures. For example, the lowest overall accuracy of CNN classifiers for an imbalanced dataset with an input image size of 128 is 90.29%. In turn, the classification rate after applying augmentation to different sizes of input images of  $(128 \times 128$  and  $64 \times 64)$  are 98.75% and 98.62%, respectively. However, the CNN-based classification method's overall accuracy is with a balanced dataset, and a  $128 \times 128$  image size is 98.75%, with an improvement of 0.13% higher than the classification result with a  $64 \times 64$  image size.

Table 4. 7 Accuracy of different breast density classification in terms of their best results.

	Size of Input Images	Class	Predicted Label by CNN				
			I	II	III	IV	
Ground truth	Imbalanced dataset (410 images)	128 × 128	I	1.0	0.0	0.0	0.0
		II	0.03	0.93	0.03	0.0	
		III	0.0	0.12	0.77	0.12	
		IV	0.0	0.0	0.33	0.67	
	64 × 64	I	0.98	0.0	0.0	0.0	
		II	0.0	0.96	0.04	0.0	
		III	0.0	0.0	1.0	0.0	
		IV	0.0	0.0	0.57	0.43	
Balanced dataset (3192 images)	128 × 128	I	1.0	0.0	0.0	0.0	
		II	0.0	0.97	0.03	0.0	
		III	0.05	0.0	0.95	0.0	
		IV	0.0	0.0	0.0	1.0	
	64 × 64	I	1.0	0.0	0.0	0.0	
		II	0.07	0.93	0.0	0.0	
		III	0.0	0.0	0.97	0.03	
		IV	0.0	0.0	0.0	1.0	

The confusion matrix of the CNN-based classification method with different input images is shown in Table 4.7. As shown, by using the balanced dataset (3192 images) with an image size of  $128 \times 128$ , the CNN classifier can correctly predict class IV with 100% accuracy as class-4 contains the high dense masks. In contrast, when the image size is changed to  $64 \times 64$ , the CNN classifier can adequately predict class I and class IV with an accuracy of 100% and class III with an accuracy of 98%. Thus, the CNN-based classification method results show how data augmentation and constructing balanced

datasets can improve the overall classification accuracy. Besides, some important objectives, such as minimizing the complexity (for example, in the LIBRA handcrafted method, they combined 86 features, in turn, in the FCN technique (Lee & Nishikawa, 2018c), seven convolution layers were used, which is more complicated than our network structure with only four layers), maximizing classification accuracy, maximizing true-positive rate, and minimizing false-positive rates are achieved in these classification approaches.

Table 4.8 Confusion matrix of breast density by CNN-based classification method.

Study	Year	Method	No. of Images	No. of Density Category	Accuracy (%)
<b>Volpara software (Highnam et al., 2010)</b>	2010	Hand-crafted	2217	Dense and Fatty	94.0
<b>LIBRA software (Keller et al., 2012a)</b>	2012	Hand-crafted	324	4 Classes	81.0
<b>Lehman et al. (Lehman et al., 2018)</b>	2018	Deep Learning	41479	Dense and non-dense	87.0
<b>Lee and Nishikawa (Lee &amp; Nishikawa, 2018a)</b>	2018	Deep Learning	455	4 Classes	85.0
<b>Mohamed et al. (Mohamed, Berg, et al., 2018)</b>	2018	Deep Learning	925	BI-RADS II and III	94.0
<b>Dubrovina et al. (Dubrovina et al., 2018)</b>	2018	Deep Learning	40	4 Classes	80.0
<b>Gadomkar et al. (Gandomkar, Suleiman, Demchig, Brennan, &amp; McEntee, 2019b)</b>	2019	Deep Learning	3813	Dense and Fatty	92.0
<b>Our proposed CNN</b>	2020	Deep Learning	410	4 Classes	98.75

For a quantitative correlation, in terms of accuracy, the performance of our proposed algorithm was compared against state-of-the-art breast density classification. A summary of representative studies can be found in Table 4.8. The accuracy, classification

method, number of datasets, and number of density categories are represented in Table 4.8. As shown, the performance of the proposed technique outperformed the state of arts with an overall accuracy of 98.75%.

## 4.5. Conclusions

Breast density can undoubtedly affect the accuracy of routine breast cancer detection methods, such as screening mammography. Therefore, it would be a breast cancer diagnostic dilemma for women with dense breast tissue (approximately 50 percent of women) (Kerlikowske et al., 2015). This chapter aims to develop an innovative and accurate method to segment and classify breast density based on the BI-RADs standard. Traditional breast density segmentation and classification methods are cumbersome tasks and have a high possibility of false positives. The efficacy of a fully automated algorithm for breast density segmentation and classification in digital mammography is proposed and substantiated by presenting three versions of cGAN networks for segmentation and two different classification methods. In our experiments, mammograms of 115 patients (410 images) from the INbreast dataset were used. With the breast density segmentation task, our method achieved an accuracy, Dice coefficient, Jaccard index of 98%, 88%, and 78%,

respectively. In addition, our method obtained precision, sensitivity, and specificity of 97.85%, 97.85%, and 99.28%, respectively, with the density classification task. Furthermore, a strong correlation between the computerized algorithm's output and the radiologist's estimated breast density can be obtained. This observation justifies that the proposed methods in this chapter have a strong positive relationship with the radiologist manual classification and are competitive with reported correlation coefficients from the literature, e.g., 0.63 (Kallenberg et al., 2016), 0.70 (Mohamed, Berg, et al., 2018), 0.85

(Li, Wei, Chan, Helvie, Roubidoux, Lu, & Zhou, 2018), and 0.85 (Lee & Nishikawa, 2018b). The most notable limitation of this chapter is that only one dataset is used. For future developments, more datasets need to be utilized; however, this dataset's ground-truth is prepared by doctors, experts, and radiologists of the Hospital Universitario Sant-Joan de Reus, Spain, via developing our GUI in MATLAB to help the radiologists to annotate the images. It is believed that artificial intelligence is capable of surpassing human experts in breast density prediction. The future work of this research is to transpose our fully automated PD% estimation techniques into the robust computer-aided breast density analyzer appraisal tool for use in clinical practice.

## **Analyzing the temporal mammographic density changes after treatment**

*If there is a recurrence, breast cancer will most likely come back within the first three years after you have finished treatment. So, it is crucial to follow up on your mammogram control and be sensitive to any changes. Mammographic density has been strongly correlated with an increased risk of breast cancer. Therefore, this chapter analyzes the three-year follow-up evolutionary of mammogram's breast density changes and the density changes in the same part of the breast where the mass was located before treatment. Breast density estimation with visual assessment is still challenging due to faint contrast and significant variations in background fatty tissues in mammograms. Hence, we applied machine learning methods with promising outcomes, leading doctors to find any changes in the breast density after treatment. Our algorithm used a cGAN network with Dice loss function, a deep learning framework for image segmentation, to segment the breast and the mass region on mammographic images. We curated an INbreast dataset and our private hospital follow-up dataset from Spain to assess its clinical setting performance and engineering performance. Four different views examine to find the analyses of the density changes after treatment. The reference density value for these images was obtained based on the consensus of two radiologists. In addition, we try to find a correlation between breast density and breast cancer subtypes, RE, RP, and Her.2.nue, which are not that significant before surgery and in control 1, but in the second control, we reach some correlation. In summary, this chapter's motivations are to identify high-risk individuals who would be potential candidates for breast cancer after treatment.*

## 5.1. Introduction

The World Health Organization (WHO)<sup>4</sup> has reported that breast cancer is the most common type of cancer in women worldwide, with 685 000 deaths globally in 2020. According to the Susan G. Komen® organization<sup>5</sup>, 7 to 11 percent of women experience a local recurrence on average. It means, after months or years of their initial treatment, breast cancer recurrence (BCR) may occur. Cancer may come back in the same place as original cancer (local recurrence), or it may spread to other areas of your body (distant recurrence). Recurrence is always possible, and BC can recur at any time, but most local recurrence occurs within the first five years after diagnosis (Colleoni et al., 2016). You cannot guarantee that your cancer will not come back; Since preventing the recurrence of breast cancer (BC) is crucial, you have to follow up your mammographic sessions after treatment to distinguish the BC risk at earlier stages when treatment could be easier and more successful. For example, Dr. Louise Eriksson and colleagues from the Karolinska Institute in Stockholm (Eriksson, Czene, Rosenberg, Humphreys, & Hall, 2013) reported that if your breast density is 25 percent or more, you have an almost two-fold increased risk of local recurrence than women with a less than 25 percent breast density. Besides, he recommended that women with dense breasts be followed more frequently or for a more extended period to spot any local recurrence quickly. Breast density (BD) is a critical BC risk factor (Valerie A. McCormack & Dos Santos Silva, 2006; Titus-Ernstoff et al., 2006; A. T. Wang, Vachon, Brandt, & Ghosh, 2014). Some studies have demonstrated that breast density has been correlated with an increased risk of local and locoregional recurrence of BC (Eriksson et al., 2013; Park et al., 2009). BD indicates

<sup>4</sup> <https://www.who.int/>

<sup>5</sup> <https://www.komen.org>

the relative amount of fibroglandular tissue, which is white in mammograms, and is calculated by dividing the dense area by the area of the whole breast (dense and non-dense tissue included). In multiple studies, longitudinal changes in breast density have been suggested to predict BC risk (Eng et al., 2014; Kanbayti, Rae, McEntee, & Ekpo, 2021; Lokate, Stellato, Veldhuis, Peeters, & Van Gils, 2013; Maskarinec, Pagano, Lurie, & Kolonel, 2006; Celine M. Vachon et al., 2007). Mammograms can be compared between temporally separated mammograms of the same breast. Radiologists use the same approach to evaluate how a suspicious region has evolved. Commonly Four different techniques (Personal, 1996) can be used for the Comparison of medical images:

- The temporal analysis compares the mammograms of the same breast at different screening examinations.
- The bilateral analysis evaluates the mammograms of the same woman's right and left breasts.
- The ipsilateral analysis is used to evaluate the craniocaudal (CC) and mediolateral oblique (MLO) views of the same breast.
- In addition, a comparison between different breast image modalities (e.g., X-ray with magnetic resonance imaging (MRI)) of the same woman is possible.

Moreover, medical image registration is a common and crucial image analysis task for comparing medical images in various biomedical image applications. It provides the ability to align one medical image with another geometrically and is essential for all imaging applications that compare datasets across subjects, imaging modalities, or time. However, very few publications can be found in the literature that discusses/address machine learning issues to the author's best knowledge. As density is a significant risk factor in the recurrence



of BC after treatment (Maskarinec, Woolcott, & Kolonel, 2010), analysis of temporal breast density changes following treatments is highly beneficial to support the doctor's decisions. On this basis, we presented a machine learning approach by analyzing the combination of whole breast density and mass region density changes by three years follow-up study to recognize and predict (who has) the potential of breast cancer recurrence; therefore, the doctors can choose a suitable treatment in an early stage. Habel (Habel et al., 2004) demonstrated that women with high breast density tend to have a higher risk of local BCR than women with low breast density. Another study (Román et al., 2019), has illustrated that changes in mammographic density over time in women could help identify the women with a higher risk of BC.

Additionally, the study of Huang (Huang et al., 2016) proved that Patients with heterogeneously dense (50–75% density) and extremely dense (>75% density) breasts had an increased risk of locoregional recurrence (hazard ratios 3.1 and 5.7, 95% confidence intervals 1.1–9.8 and 1.2–34.9,  $p = 0.043$  and  $0.048$ , respectively) than did women with less dense breasts. As well as this, a majority of the studies that assessed breast density changes compared breast density measured from post-treatment mammograms to that measured from mammograms acquired two or more years before diagnosis (screening mammograms) (Engmann et al., 2017; Knight et al., 2018; Nyante et al., 2015, 2016). Besides, some studies investigated (J. Kim et al., 2012; Ko et al., 2013; Nyante et al., 2015) decreasing the breast density during the follow-up BC treatment reduces the risk of recurrence and death from the disease.

Some other studies (Cuzick, Warwick, Pinney, Warren, & Duffy, 2004; J. Kim et al., 2012; Nyante et al., 2015) investigated density reductions with tamoxifen therapy of primary

breast cancer recurrence and mortality. In addition, it has been shown that women who experience a 10% reduction or more in breast density following adjuvant endocrine treatments have better treatment outcomes (Anampa, Makower, & Sparano, 2015; J. Kim et al., 2012; Ko et al., 2013; Nyante et al., 2016).

Findings from this chapter should provide an understanding of breast density trends and factors influencing breast density after BC treatment. It could also help to identify the category of women that breast density can be used to monitor the effectiveness of adjuvant therapy in reducing BC recurrence and mortality (Kanbayti et al., 2021). Finally, evaluate these changes within a maximum of 3 years (Eriksson et al., 2018; Nyante et al., 2016; C. M. Vachon et al., 2007).

In this chapter, we describe the novel computer method of representing breast density changes after treatment to assist oncologists in identifying high-risk individuals who would be potential candidates for breast cancer recurrence after treatment. We analyzed and monitored the combination of the breast tissue density (global breast density), and mass region density (local breast density) changes for three years after treatment on mammograms. Our ability to predict the suspicious cases of breast cancer risk after treatment is highly likely to improve in the future; we provide a tested method that could be used to predict breast cancer recurrence risk. In addition, the correlation between the breast density and molecular subtypes of breast cancer, estrogen receptor (ER), progesterone receptor (PR), Ki67 antigen (Ki-67), and human epidermal growth factor receptor-2 (HER2) is also provided. Further sections of this literature are as follows: Section 2 contains the methodology and the search process, the results and discussion are given in section 3, and section 4 shows the conclusion.

## 5.2. Methodology

The overview of research methodology for analyzing the temporal breast density changes and predicting the risky case after treatment is shown in Figure 5.1, consisting of three stages. In the first stage, whole-breast density was analyzed for three years of follow-up controls, which led to a score1. In the second stage, after extracting the mass region in three years follow-up controls by registration method, the density of the mass region is analyzed, which leads to score2. Then, by calculating the multiple of Score 1 and 2, we found the threshold for non- risky control and risky cases in each follow-up mammogram. Which Score 1 is considered a risk score in the whole breast and score 2 is a risk demonstrated in the mass region,  $n$  is a control number, and  $i$  is a patient ID. A study of the correlation between the breast density and molecular subtypes of breast cancer, estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor-2 (HER2) is also provided. All steps are described in detail in the following sections.

### Implementation details

The proposed method was applied using Python with PyTorch <sup>6</sup>, going on a 64-bit Ubuntu operating system utilizing a 3.4 GHz Intel Core-i7 with 16 GB of RAM and Nvidia GTX 1070 GPU with 8 GB of video RAM.

<sup>6</sup> <https://pytorch.org>

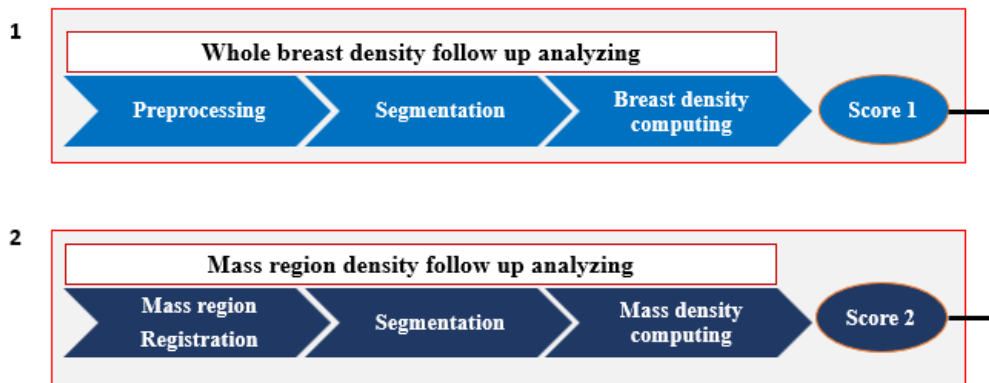


Figure 5. 1 Methodology of breast cancer risk based on temporal breast density analyses

### 5.2.1. Global Breast density follow up analyzing

The proposed whole breast density follow-up analysis method shown in Figure 5.1 comprises three steps.

- In the first stage, mammograms are pre-processed, consisting of three steps: Resizing the generated mask images to the same resolution as the input mammography (512 × 512 pixels), then removing the background and the pectoral muscle.
- Next, the conditional generative adversarial network (cGAN) with dice loss function used in our previous work (Saffari et al., 2020) is applied to segment the dense tissue.
- Finally, computing the percentage of breast density by dividing the dense tissues' area by the total breast area (the result is between 0 to 100).

As shown in Figure 5.2, the conditional generative adversarial network, or cGAN for short, is a conditional type of the GAN. The cGAN network consists of two networks. The first network is an encoder-decoder network, called a generator, that maps the input mammogram to a segmented

image, including the dense areas. In turn, the discriminator learns a loss function to enforce the generator to produce an output similar to the ground truth.

In our model, the generated network takes a mammographic image and tries to generate a mask image of the areas related to dense tissues (i.e., 0 for non-dense pixels including the background pixels and 1 for dense tissues pixels). The generator network then generates data latterly fed into a discriminator network. Finally, the discriminator learns a loss function to train this mapping by comparing the ground-truth and the predicted output but observing the input image as a condition to improve the network optimization. The non-zero pixels in the generated mask counted for signifying the dense tissue area. Then, the breast density in the input images is calculated by Computing the ratio between the area of dense tissues and the area of the breast region. We applied these steps in each patient's ( $i = \text{patient's ID}$ ) four mammogram views  $MLO_{Right} = V_1$ ,  $MLO_{Left} = V_2$ ,  $CC_{Right} = V_3$ ,  $CC_{Left} = V_4$ . Finally, to find the risky case in each control, we calculate the (score 1  $\times$  score2) of each follow-up mammogram ( $CTRL_n$ ) of the same patients ( $i = \text{patient's ID}$ ) presented in (5.6).

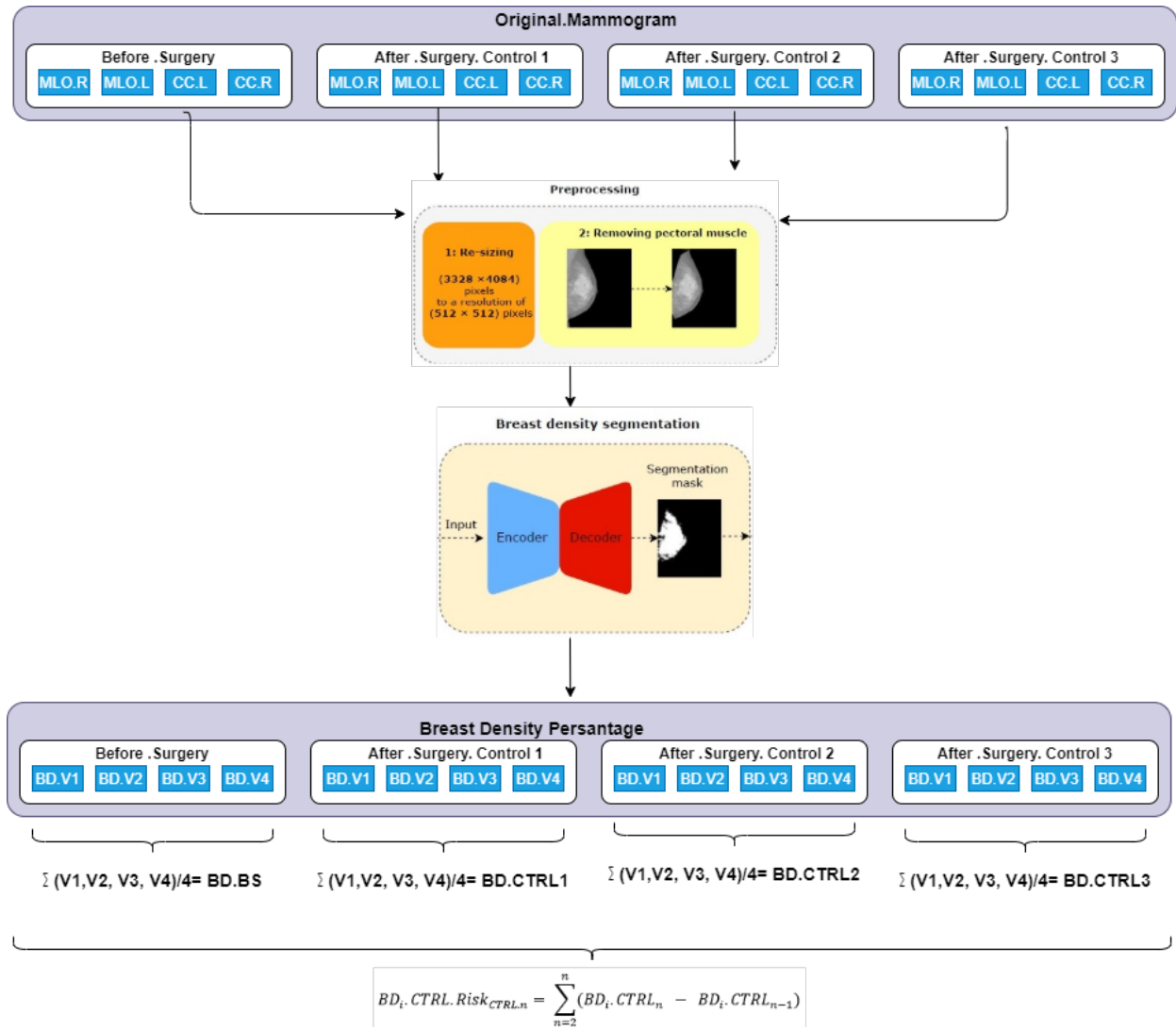


Figure 5. 2 Analysing the longitudinal of Whole breast density in three years (V1 is MLO-R, V2 is MLO-L, V3 is CC-R, V4 is CC-L)

As shown in Figure 5.2, to find out the breast density changes during these three follow-up years first, we calculated the global breast density in different views of each control ( $i$  = patient's ID and  $n$  view's number) then, the average of 4 views considered as a breast density patient's ( $i$ ) of the first control after surgery, second control after surgery and third control after surgery, that illustrated in (5.1), in the 3rd step which is shown in the in (5.2), we calculated the difference of breast density control ( $n$ ) and ( $n-1$ ) to find out the risky case of each patient's  $i$ . Then, to distinguish between risky and non-risky cases for each

patient, we defined the threshold based on each control's estimated global breast density.

The threshold of 20 ( *non – risky*  $\leq 20$  and *risky*  $\geq 20$  ) is considered as a boundary between risky and non-risky, which is presented in (5.2). As illustrated in (5.2), when we have a risky case,  $Score\ 1_{CTRL.n_i} = 1$ .

$$BD_{i_{CTRL.n}} = \frac{1}{4} \sum (BD_{i_{v_1}} + BD_{i_{v_2}} + BD_{i_{v_3}} + BD_{i_{v_4}}) \quad (5.1)$$

$$BD_{i.CTRL.Risk_{CTRL.n}} = \sum_{n=2}^n (BD_{i.CTRL_n} - BD_{i.CTRL_{n-1}}) \quad (5.2)$$

$$\begin{cases} \text{if } BD_{i.CTRL.Risk_{CTRL.n}} \geq \text{threshold} & \text{then } Score\ 1_{CTRL.n_i} = 1 \\ \text{Else} & Score\ 1_{CTRL.n_i} = 0 \end{cases}$$

### 5.2.2. Local density follows up analyzing

Mammogram registration helps doctors analyze and visualize the comparison of the current mammogram with the previous one. Registration (alignment) method is required for mammograms comparison. As shown in Figure 5.3, the registration method was applied to analyze the density changes in the local region of the temporal mammogram images. The temporal mammogram images registration is used to extract the same location of the mass region from the original mammograms in Control.1, Control.2, and Control.3 mammograms. The ratio between the area of the mass region extracted and the non-zero pixels in the cropped mass region is calculated for measuring the mass region density. These processes are applied in each patient's mammograms with mass. The average of different views value was considered a mass density of each year illustrated in (5.3). Then, to distinguish the risky local case of each control of patient's *i*, we calculated the difference of mass region's density percentage of control (*n*) and (*n-1*) of all patients

$i$ , which is shown in (5.4). Finally, to distinguish between risky and non-risky cases for each patient, we defined the threshold based on each control's estimated local breast density. The threshold of 15 (  $non - risky \leq 15$  and  $risky \geq 15$  ) is considered as a boundary between risky and non-risky, which is presented in (5.4). As you can see when we have a risky local case,  $Score 2_{CTRL.n_i} = 1$ .

$$MD_{i_{CTRL.n}} = \frac{1}{4} \sum (MD_{i_{v_1}} + MD_{i_{v_2}} + MD_{i_{v_3}} + MD_{i_{v_4}}) \quad (5.3)$$

$$MD_i.CTRL.Risk_{CTRL.n} = \sum_{n=2}^n (MD_i.CTRL_n - MD_i.CTRL_{n-1}) \quad (5.4)$$

$$\begin{cases} \text{if } MD_i.CTRL.Risk_{CTRL.n} \geq \text{threshold} & \text{then } Score 2_{CTRL.n_i} = 1 \\ \text{Else} & Score 2_{CTRL.n_i} = 0 \end{cases}$$

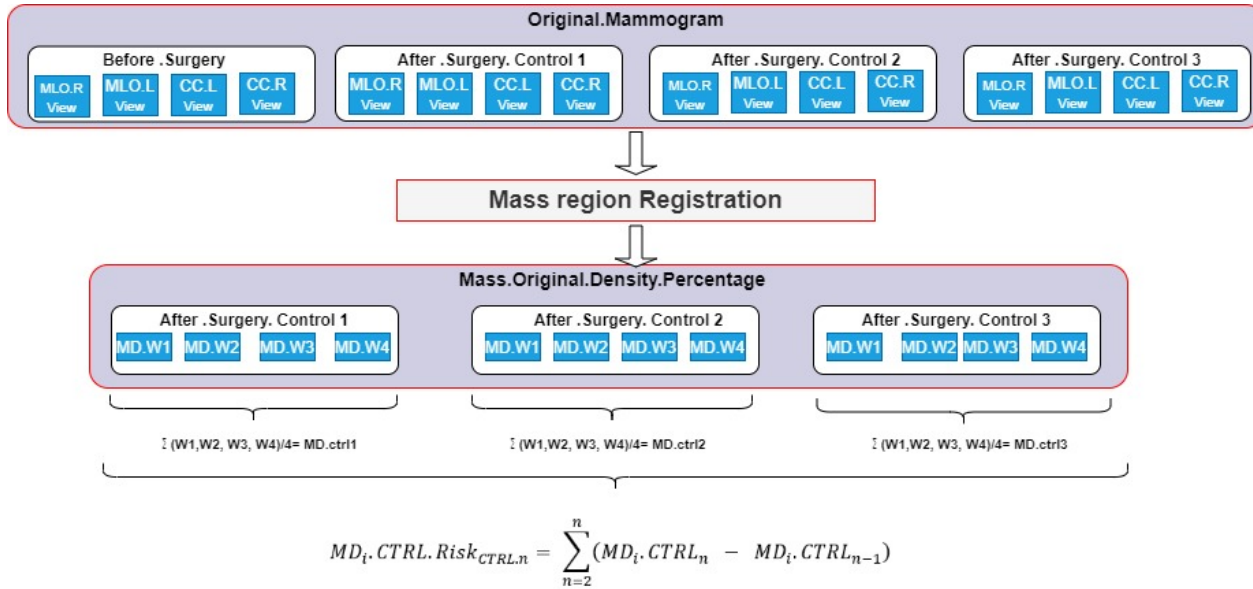


Figure 5.3 Analysing the longitudinal of mass region density in three years (W1 is MLO-R, W2 is MLO-L, W3 is CC-R, W4 is CC-L)



### 5.2.3. Breast density changes overall index

To determine the risky case, the multiple of whole breast density (global) changes, score1, and mass region density (local) change, score 2, was determined as an overall risky result. It means if the difference value of global and local breast density percentage both of control ( $n$ ) and ( $n-1$ ) of patient  $i$  is more than the threshold, control ( $n$ ) of patient  $i$  considers risky cases, which is presented in (5.5).

$$\begin{cases} \text{if } \text{Score 2 CTRL. } n_i \times \text{Score 1 CTRL. } n_i = 1 & \text{is a risky case} \\ \text{Else} & \text{is not risky} \end{cases} \quad (5.5)$$

## 5.3. Experimental results

### 5.3.1. Dataset

We have evaluated the performance of proposed models on **two** mammography datasets:

#### **INbreast dataset**<sup>7</sup>

INbreast dataset is a publicly available 2-dimensional (2D) database and includes MLO and CC mammographic images of 115 patients (410 mammograms). Every patient has four mammographic images (which consists of MLO-right, MLO-left, CC-right, and CC-left). It has a ground truth for mass location, mass type, and breast density classification label. Breast density classification in INbreast was prepared based on breast imaging reporting and data method (BI-RADS) standard (BI-RADS I, BI-RADS II, BI-RADS III, and BI-RADS IV).

<sup>7</sup> [http://medicalresearch.inescporto.pt/breastcancer/index.php/Get\\_INbreast\\_Database/](http://medicalresearch.inescporto.pt/breastcancer/index.php/Get_INbreast_Database/)

The image size of the mammogram is  $3328 \times 4084$  or  $2560 \times 3328$  pixels. The INbreast dataset does not have the ground truth binary masks for the breast density segmentation (We have the label of each image corresponding to the breast density classification ((BI-RADS I, BI-RADS II, BI-RADS III, and BI-RADS IV). In this chapter, two radiologists collaborated with the research team to provide the whole dataset's ground truth by segmenting the dense tissues in mammograms. We train the model by 410 mammograms of the INbreast dataset, which has a density classification label. The proposed cGAN segmentation model is trained by the INbreast dataset.

### **Hospital Sant Joan de Reus dataset**

Our private dataset includes 50 patients with their respective three-year follow-up mammograms from 4 different views after surgery. The suggested cGAN segmentation model (Saffari et al., 2020) was tested on this dataset. The duty of confidentiality and security measures fully complied with the current legislation protecting personal data (Article 7.1 of the Organic Law 15/1999, 13th of December).

## **5.3.2. Result of breast density segmentation analysis**

To investigate the breast density changes after treatment, we analyze the whole breast density and mass region density changes to detect the suspicious patinas; First, we segmented the mammogram images; then calculated the average density of four views (MLO-right, MLO-left, CC-right, and CC-left). The sample of segmented mammograms of one Patient in four different views shows in Figure 5. 4.

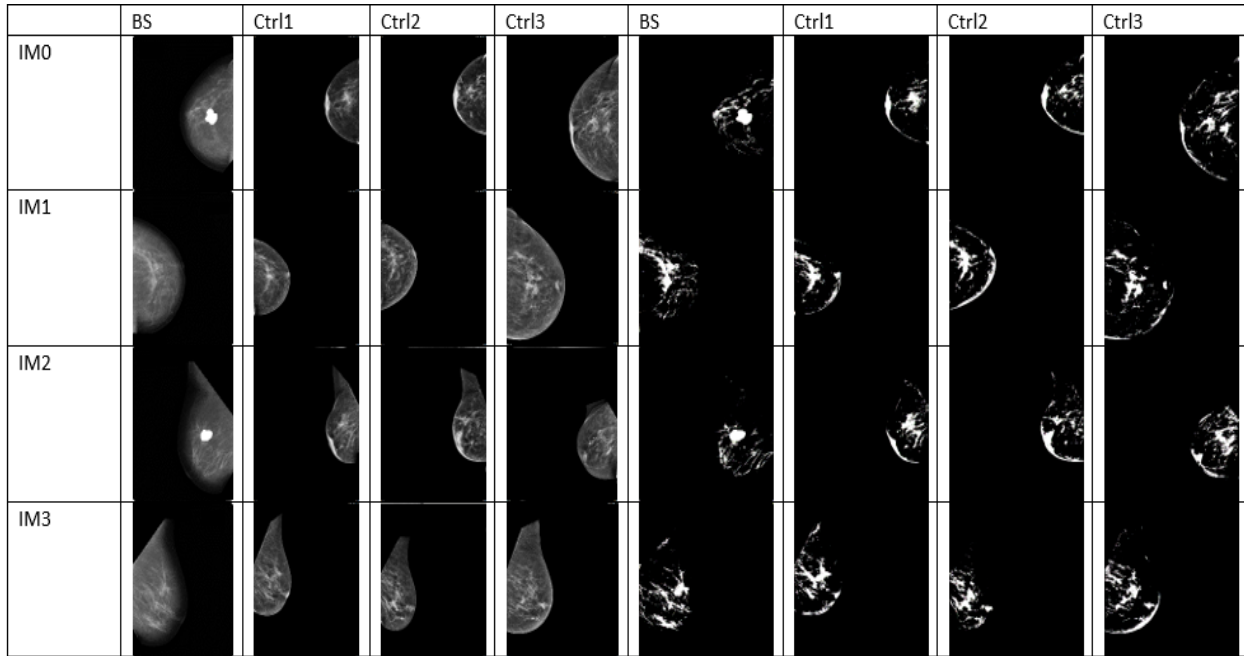


Figure 5. 4 samples of temporal breast density segmentation in three years

The geometric transformation registration and Intensity-based image registration method were investigated to extract the exact location of the mass region from the original mammograms in different follow-up control mammograms after treatment. In Figure 5.5, you can see the global and local density analysis sample results for one Patient with two years of follow-up. As you can see, we don't have a significant density change for this Patient; in other words, this Patient's breast density does not change significantly, so it is not a risky case.

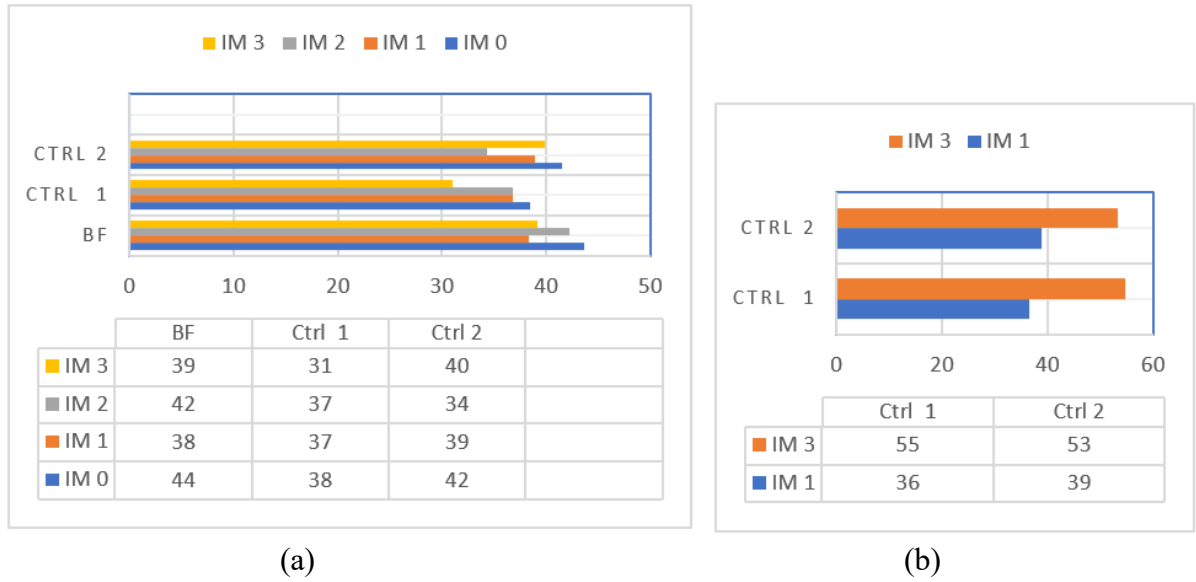


Figure 5.5 sample of global (a) and local (b) breast density analyzing for one non-risky patient with two years follow up

## 5.4. Conclusions

This chapter focused on analyzing the local and general breast density changes in temporal follow-up mammograms after treatment based on deep learning methods. First, to segmentation the tissue, we applied cGAN, which can learn the intrinsic features of the density from a relatively low number (hundreds) of training samples and then generate the corresponding image mask that selects the pixels belonging to the tissue. The purpose is to find the threshold between risky and non-risky cases. The threshold of local breast density changes is 20. If the whole breast density difference between control ( $n$ ) and ( $n-1$ ) of patient  $i$  is more than 20, we consider patient  $i$ , as a risky global case. On the other hand, if the threshold of local breast density changes of patient  $i$  be more than 15, we say that control  $n$  of patient  $i$ , in general, is a risky case and needs more analysis.

## **Concluding remarks**

*This chapter presents the most significant contributions and main conclusions of this dissertation, highlighting their significance. Furthermore, the chapter also involves proposals for future work.*

## 6.1. Thesis highlights

Breast density is a prevalent high-risk cause of breast cancer in women. Numerous studies have reported that women with high-density breasts develop breast cancer independent of their age, menopausal status, and hormone replacement therapy. Therefore, improving the screening process for women with high breast density is vital and can decrease the risk of breast cancer, so it is essential to report the breast density to guide women with high breast density for additional examinations, such as ultrasonography or MRI.

The main goal of this doctoral dissertation project was to investigate the fully automated segmentation and classification scheme for mammograms based on breast density estimation using deep learning methods to overcome this challenge. The findings presented in this doctoral dissertation is promising and show that the proposed technique can produce a clinically helpful computer-aided digital tool for breast density analysis by digital mammography.

The present thesis makes several noteworthy contributions to the literature on developing innovative algorithms based on machine learning and deep learning techniques to improve and increase the accuracy of fully automated mammograms computer-aided design (CAD). In addition, this research has several practical applications, which are presented in four chapters and summarized below.

Chapter 2 proposes a new methodology by combining the traditional texture analysis methods with deep learning approaches to classify breast tumors into malignant or benign. This method used a bio-inspired optimization algorithm (also known as a whale optimization algorithm) to learn local descriptors from the input images. Ultrasound and mammographic

images datasets with benign and malignant cases were used to assess the proposed method. This research showed the accuracy of this model on both ultrasound and mammographic images was within acceptable ranges.

In Chapter 3, a novel breast density segmentation methodology is developed based on the conditional Generative Adversarial Network. The cGAN network consists of two networks, the first network is an encoder-decoder network (also called the generator) that maps the input mammogram to a segmented image including the dense areas, and the second network is called discriminator that learns a loss function to enforce the generator to produce an output, which is like the ground-truth. Using the cGAN network helps the proposed model to train by a small set of mammographic images. Furthermore, the proposed method can segment the dense regions in mammographic images with outstanding performance compared to the standard semantic segmentation methods (FCN and SegNet).

In Chapter 4, an innovative and accurate methodology is developed for the segmentation and classification of breast density based on the BI-rads standard. Unfortunately, traditional breast density segmentation and classification methods are complex tasks and have a high possibility of false positives. To overcome this problem, the efficacy of a fully automated algorithm for breast density segmentation and classification in digital mammography is proposed and substantiated by presenting three versions of cGAN networks for segmentation and two different classification methods.

In the experiments carried out in this chapter, mammograms of 115 patients (410 images) from the INbreast dataset were utilized. The breast density segmentation study showed that the proposed cGAN with dice loss function could achieve achieved an accuracy of 98%, which is the highest accuracy compared to other loss functions.

In addition, using the proposed CNN methodology for density classification leads to an accuracy equal to 98.75%. Finally, it is noteworthy that a strong correlation between the computerized algorithm's output and the radiologist's estimated breast density can be obtained. This observation justifies that the proposed methods in this chapter have a strong positive relationship with the radiologist manual classification and are competitive with reported correlation coefficients from the literature.

For future developments, more datasets need to be utilized; however, the ground-truth of the dataset needs to be prepared by doctors, experts, and radiologists of the Hospital Universitari Sant-Joan of Reus, Spain, via developing our Graphical User Interface (GUI) in MATLAB to help the radiologists to annotate the images. It is believed that artificial intelligence can surpass human experts in breast density prediction.

One of the most important practical applications of this Ph.D. project which can emerge in the near future will be transposing the fully automated PD% estimation techniques developed in this thesis into a robust computer-aided breast density analyzer appraisal tool/software for the public use in clinical practice.

Chapter 5 analyzes the local and general breast density changes in temporal follow-up mammograms after treatment based on deep learning methods. For tissue segmentation, the cGAN technique was applied to learn the intrinsic features of the density from a relatively low number (hundreds) of training samples and then generate the corresponding image mask that selects the pixels belonging to the tissue.

The main aim of this methodology is to determine the threshold between risky and non-risky cases. If the global and local breast density difference between control “ $n$ ” and “ $n$ -



If the value of  $I_i$  is more than a threshold, patient  $i$  is considered as a risky case and requires more analysis.

## 6.2. Future research lines

The work presented in this thesis addresses the interpretation of breast density analysis. We believe this is a crucial field of AI research. Several directions of future work have been identified during this work which we couldn't investigate some of them due to time limitations. For example:

1. The medical image datasets to train the model are usually small in the number of samples. This condition decreases the generalization ability of the model, which means that it may fail to accurately process test examples that present small changes in the expected patterns. By increasing the number of patients, we can develop a more accurate and robust deep learning algorithm, increasing the efficacy of the breast density CAD systems.
2. The purpose of cancer treatments is to eliminate cancer cells. However, cancer cells are tricky. Treatments can reduce tumors so much that examinations do not detect their presence. These weakened cells can remain in the body after treatment. Over time, the cells get more robust. They start to grow and multiply again. It means even surgery to remove a cancerous tumor is not always 100% effective. Cancer may come back in the same place as original cancer (local recurrence), or it may spread to other areas of your body (distant recurrence). If there's a recurrence, breast cancer is most likely to come back within the first two years after your treatment. So it's crucial to pay attention to your follow-up,

particularly during this time. By analyzing the breast cancer recurrence cases, we can predict any phenomenons early and help the patients survive.

3. The current state of art mammogram CAD systems only accepts images without considering the patient's clinical data. Yet, in practice, the patient's clinical history and laboratory data enable physicians to interpret imaging findings in the appropriate clinical context leading to higher and more accurate diagnostics. Different data fusion techniques can be applied to combine the mammogram breast images with patients' clinical and laboratory data to achieve this goal and improve the accuracy of the diagnostics.
4. Breast cancer can come back in another part of the body(metastatic) months or years after the original diagnosis and treatment. It is estimated that 20% to 30% of women diagnosed with early-stage breast cancer will develop metastatic disease. By studying the metastatic datasets, we can find the effective biomarkers, behavior, and patterns of metastatic patients to predict the risky metastatic cases in the early stage.
5. We can add the blood test of each patient and find the correlation of breast cancer risk perdition based on breast density and the blood test result, which is the current method to detect breast cancer recurrence.

## References

---

- Abbas, Q. (2016). DeepCAD: A computer-aided diagnosis system for mammographic masses using deep invariant features. *Computers*, 5(4). <https://doi.org/10.3390/computers5040028>
- Abdel-Nasser, M., Melendez, J., Moreno, A., Omer, O. A., & Puig, D. (2017). Breast tumor classification in ultrasound images using texture analysis and super-resolution methods. *Engineering Applications of Artificial Intelligence*, 59(December 2016), 84–92. <https://doi.org/10.1016/j.engappai.2016.12.019>
- Abdel-Nasser, M., Melendez, J., Moreno, A., & Puig, D. (2016). The impact of pixel resolution, integration scale, preprocessing, and feature normalization on texture analysis for mass classification in mammograms. *International Journal of Optics*, 2016. <https://doi.org/10.1155/2016/1370259>
- Abdel-Nasser, M., Moreno, A., A. Rashwan, H., & Puig, D. (2016). Analyzing the evolution of breast tumors through flow fields and strain tensors. *Pattern Recognition Letters*. <https://doi.org/10.1016/j.patrec.2016.11.003>
- Abdel-nasser, M., Moreno, A., & Puig, D. (2016). Temporal mammogram image registration using optimized curvilinear coordinates. *Computer Methods and Programs in Biomedicine*, 7, 1–14. <https://doi.org/10.1016/j.cmpb.2016.01.019>
- Abdel-Nasser, M., Rashwan, H. A., Puig, D., & Moreno, A. (2015). Analysis of tissue abnormality and breast density in mammographic images using a uniform local directional pattern. *Expert Systems with Applications*, 42(24), 9499–9511. <https://doi.org/10.1016/j.eswa.2015.07.072>
- Abdel-Nasser, M., Saleh, A., Moreno, A., Tabalvandani, N. S., & Puig, D. (2017). Feature learning

- for breast tumour classification using bio-inspired optimization algorithms. *Frontiers in Artificial Intelligence and Applications*, 300, 106–115. <https://doi.org/10.3233/978-1-61499-806-8-106>
- Ahn, C. K., Heo, C., Jin, H., & Kim, J. H. (2017). *A novel deep learning-based approach to high accuracy breast density estimation in digital mammography*. 10134(Cc), 1013420. <https://doi.org/10.1117/12.2254264>
- Anampa, J., Makower, D., & Sparano, J. A. (2015). Progress in adjuvant chemotherapy for breast cancer: An overview. *BMC Medicine*, 13(1), 1–13. <https://doi.org/10.1186/s12916-015-0439-8>
- Arevalo, J., González, F. A., Ramos-Pollán, R., Oliveira, J. L., & Guevara Lopez, M. A. (2016). Representation learning for mammography mass lesion classification with convolutional neural networks. *Computer Methods and Programs in Biomedicine*, 127(December 2015), 248–257. <https://doi.org/10.1016/j.cmpb.2015.12.014>
- Arnau Oliver, Jordi Freixenet, Robert Marti, Josep Pont, Elsa Pérez, Erika R. E. Denton, and R. Z. (2008). A Novel Breast Tissue Density Classification Methodology Arnau. *IEEE TRANSACTIONS ON INFORMATION TECHNOLOGY IN BIOMEDICINE*, 12(1), 55–65. Retrieved from <http://www.freepatentsonline.com/y2004/0024620.html>
- Astley, S. M., Harkness, E. F., Sergeant, J. C., Warwick, J., Stavrinou, P., Warren, R., ... Evans, D. G. (2018). *A comparison of five methods of measuring mammographic density: a case-control study*. 1–13. <https://doi.org/10.1186/s13058-018-0932-z>
- Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and*

*Machine Intelligence*, 39(12), 2481–2495. <https://doi.org/10.1109/TPAMI.2016.2644615>

Becker, A. S., Marcon, M., Ghafoor, S., Wurnig, M. C., Frauenfelder, T., & Boss, A. (2017). Deep learning in mammography diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer. *Investigative Radiology*, 52(7), 434–440. <https://doi.org/10.1097/RLI.0000000000000358>

Boyd, N. F., Lockwood, G. A., Byng, J. W., Tritchler, D. L., & Yaffe, M. J. (1998). Mammographic densities and breast cancer risk. *Cancer Epidemiology Biomarkers and Prevention*, 7(12), 1133–1144.

Boyd, N. F., Martin, L. J., Bronskill, M., Yaffe, M. J., Duric, N., & Minkin, S. (2010). Breast tissue composition and susceptibility to breast cancer. *Journal of the National Cancer Institute*, 102(16), 1224–1237. <https://doi.org/10.1093/jnci/djq239>

Breast Cancer U.S. (2016). ‘Breast Cancer Statistics,’ Available online: [http://www.breastcancer.org/symptoms/understand\\_bc/statistics](http://www.breastcancer.org/symptoms/understand_bc/statistics).

*BREAST PATTERNS AS INDEX OF BREAST*. (1976). 1130–1139.

Byng J W, Boyd N F, Fishell E, Jong R A, Y. M. J. (1994). Physics in Medicine & Biology Related content The quantitative analysis of mammographic densities The quantitative analysis of mammographic densities. *Physics in Medicine and Biology*, 1629–1638.

Cancer Research UK. (2014). ‘Breast cancer statistics Available online’ <http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer>.

Chan, H. P., & Helvie, M. A. (2019). Deep learning for mammographic breast density assessment

and beyond. *Radiology*, 290(1), 59–60. <https://doi.org/10.1148/radiol.2018182116>

Chen, Y., Wang, Y., & Yang, B. (2006). Evolving hierarchical RBF neural networks for breast cancer detection. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 4234 LNCS, 137–144. [https://doi.org/10.1007/11893295\\_16](https://doi.org/10.1007/11893295_16)

Ciatto, S., Bernardi, D., Calabrese, M., Durando, M., Gentilini, M. A., Mariscotti, G., ... Houssami, N. (2012). A first evaluation of breast radiological density assessment by QUANTRA software as compared to visual classification. *Breast*, 21(4), 503–506. <https://doi.org/10.1016/j.breast.2012.01.005>

Ciritsis, A., Rossi, C., De Martini, I. V., Eberhard, M., Marcon, M., Becker, A. S., ... Boss, A. (2019). Determination of mammographic breast density using a deep convolutional neural network. *British Journal of Radiology*, 92(1093). <https://doi.org/10.1259/bjr.20180691>

Colleoni, M., Sun, Z., Price, K. N., Karlsson, P., Forbes, J. F., Thürlimann, B., ... Goldhirsch, A. (2016). Annual hazard rates of recurrence for breast cancer during 24 years of follow-up: Results from the international breast cancer study group trials I to V. *Journal of Clinical Oncology*, 34(9), 927–935. <https://doi.org/10.1200/JCO.2015.62.3504>

Cuzick, J., Warwick, J., Pinney, E., Warren, R. M. L., & Duffy, S. W. (2004). Tamoxifen and breast density in women at increased risk of breast cancer. *Journal of the National Cancer Institute*, 96(8), 621–628. <https://doi.org/10.1093/jnci/djh106>

Dalmış, M. U., Litjens, G., Holland, K., Setio, A., Mann, R., Karssemeijer, N., & Gubern-Mérida, A. (2017). Using deep learning to segment breast and fibroglandular tissue in MRI volumes: *Medical Physics*, 44(2), 533–546. <https://doi.org/10.1002/mp.12079>

- Dhungel, N., Carneiro, G., & Bradley, A. P. (2017). A deep learning approach for the analysis of masses in mammograms with minimal user intervention. *Medical Image Analysis*, *37*, 114–128. <https://doi.org/10.1016/j.media.2017.01.009>
- Dubrovina, A., Kisilev, P., Ginsburg, B., Hashoul, S., & Kimmel, R. (2018). Computational mammography using deep neural networks. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization*, *6*(3), 243–247. <https://doi.org/10.1080/21681163.2015.1131197>
- Eng, A., Gallant, Z., Shepherd, J., McCormack, V., Li, J., Dowsett, M., ... Dos-Santos-Silva, I. (2014). Digital mammographic density and breast cancer risk: a case-control study of six alternative density assessment methods. *Breast Cancer Research: BCR*, *16*(5), 439. <https://doi.org/10.1186/s13058-014-0439-1>
- Engmann, N. J., Scott, C. G., Jensen, M. R., Ma, L., Brandt, K. R., Mahmoudzadeh, A. P., ... Kerlikowske, K. (2017). Longitudinal changes in volumetric breast density with tamoxifen and aromatase inhibitors. *Cancer Epidemiology Biomarkers and Prevention*, *26*(6), 930–937. <https://doi.org/10.1158/1055-9965.EPI-16-0882>
- Eriksson, L., Czene, K., Rosenberg, L., Humphreys, K., & Hall, P. (2013). Possible influence of mammographic density on local and locoregional recurrence of breast cancer. *Breast Cancer Research*, *15*(4). <https://doi.org/10.1186/bcr3450>
- Eriksson, L., He, W., Eriksson, M., Humphreys, K., Bergh, J., Hall, P., & Czene, K. (2018). Adjuvant Therapy and Mammographic Density Changes in Women With Breast Cancer. *JNCI Cancer Spectrum*, *2*(4), 1–8. <https://doi.org/10.1093/jncics/pky071>
- Figueredo, A. J., & Wolf, P. S. A. (2009). Assortative pairing and life history strategy. *Human*

*Nature*, 20(3), 317–330. <https://doi.org/10.1007/s12110-009-9068-2>

Gamdonkar, Z., Tay, K., Ryder, W., Brennan, P. C., & Mello-Thoms, C. (2015). *iDensity: an automatic Gabor filter-based algorithm for breast density assessment*. 9416, 941607. <https://doi.org/10.1117/12.2083149>

Gandomkar, Z., Suleiman, M. E., Demchig, D., Brennan, P. C., & McEntee, M. F. (2019a). *BI-RADS density categorization using deep neural networks*. <https://doi.org/10.1117/12.2513185>

Gandomkar, Z., Suleiman, M. E., Demchig, D., Brennan, P. C., & McEntee, M. F. (2019b). *BI-RADS density categorization using deep neural networks*. (March 2019), 22. <https://doi.org/10.1117/12.2513185>

Gram, I. T., Funkhouser, E., & Tabár, L. (1997). The Tabar classification of mammographic parenchymal patterns. *European Journal of Radiology*, 24(2), 131–136. [https://doi.org/10.1016/S0720-048X\(96\)01138-2](https://doi.org/10.1016/S0720-048X(96)01138-2)

Habel, L. A., Dignam, J. J., Land, S. R., Salane, M., Capra, A. M., & Julian, T. B. (2004). Mammographic density and breast cancer after ductal carcinoma in situ. *Journal of the National Cancer Institute*, 96(19), 1467–1472. <https://doi.org/10.1093/jnci/djh260>

Harvey, J. A., & Bovbjerg, V. E. (2004). *Radiology Abbreviations : Quantitative Assessment of Mammographic Breast Density : Relationship with Breast Cancer Risk 1 COMPOSITION OVER TIME : 29–41*.

He, W., Denton, E. R. E., Stafford, K., & Zwiggelaar, R. (2011). Mammographic image segmentation and risk classification based on mammographic parenchymal patterns and



geometric moments. *Biomedical Signal Processing and Control*, 6(3), 321–329.  
<https://doi.org/10.1016/j.bspc.2011.03.008>

Highnam, R., Brady, M., Yaffe, M. J., Karssemeijer, N., & Harvey, J. (2010). Robust breast composition measurement - Volpara™. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6136 LNCS, 342–349. [https://doi.org/10.1007/978-3-642-13666-5\\_46](https://doi.org/10.1007/978-3-642-13666-5_46)

Huang, Y.-S., Chen, J. L.-Y., Huang, C.-S., Kuo, S.-H., Jaw, F.-S., Tseng, Y.-H., ... Chang, Y.-C. (2016). High mammographic breast density predicts locoregional recurrence after modified radical mastectomy for invasive breast cancer: a case-control study. *Breast Cancer Research*, 18(1), 120. <https://doi.org/10.1186/s13058-016-0784-3>

Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017a). Image-to-image translation with conditional adversarial networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua*, 5967–5976. <https://doi.org/10.1109/CVPR.2017.632>

Isola, P., Zhu, J. Y., Zhou, T., & Efros, A. A. (2017b). Image-to-image translation with conditional adversarial networks. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua*, 5967–5976. <https://doi.org/10.1109/CVPR.2017.632>

Jadoon, M. M., Zhang, Q., Haq, I. U., Butt, S., & Jadoon, A. (2017). Three-Class Mammogram Classification Based on Descriptive CNN Features. *BioMed Research International*, 2017. <https://doi.org/10.1155/2017/3640901>

Kallenberg, M., Petersen, K., Nielsen, M., Ng, A. Y., Diao, P., Igel, C., ... Lillholm, M. (2016). Unsupervised Deep Learning Applied to Breast Density Segmentation and Mammographic Risk Scoring. *IEEE Transactions on Medical Imaging*, 35(5), 1322–1331.

<https://doi.org/10.1109/TMI.2016.2532122>

Kanbayti, I. H., Rae, W. I. D., McEntee, M. F., & Ekpo, E. U. (2021). Mammographic density changes following BC treatment. *Clinical Imaging*, 76(December 2020), 88–97.

<https://doi.org/10.1016/j.clinimag.2021.01.002>

Keller, B. M., Nathan, D. L., Wang, Y., Zheng, Y., Gee, J. C., Conant, E. F., & Kontos, D. (n.d.).

*Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation.*

Keller, B. M., Nathan, D. L., Wang, Y., Zheng, Y., Gee, J. C., Conant, E. F., & Kontos, D. (2012a).

Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation.

*Medical Physics*, 39(8), 4903–4917. <https://doi.org/10.1118/1.4736530>

Keller, B. M., Nathan, D. L., Wang, Y., Zheng, Y., Gee, J. C., Conant, E. F., & Kontos, D. (2012b).

*Estimation of breast percent density in raw and processed full field digital mammography images via adaptive fuzzy c-means clustering and support vector machine segmentation.*

39(August), 4903–4917.

Kerlikowske, K., Zhu, W., Tosteson, A. N. A., Sprague, B. L., Tice, J. A., Lehman, C. D., &

Miglioretti, D. L. (2015). Identifying women with dense breasts at high risk for interval cancer a cohort study. *Annals of Internal Medicine*. <https://doi.org/10.7326/M14-1465>

Kim, J., Han, W., Moon, H. G., Ahn, S. K., Shin, H. C., You, J. M., ... Noh, D. Y. (2012).

Correction: Breast density change as a predictive surrogate for response to adjuvant endocrine therapy in hormone receptor positive breast cancer. *Breast Cancer Research*, 14(6), 0–7.

<https://doi.org/10.1186/bcr3353>

- Kim, Y., Kim, C., & Kim, J.-H. (2010). Automated estimation of breast density on mammogram using combined information of histogram statistics and boundary gradients. *Medical Imaging 2010: Computer-Aided Diagnosis*, 7624(March 2010), 76242F. <https://doi.org/10.1117/12.844083>
- Kingma, D. P., & Ba, J. (2014). *Adam: A Method for Stochastic Optimization*. 1–15. <https://doi.org/http://doi.acm.org.ezproxy.lib.ucf.edu/10.1145/1830483.1830503>
- Knight, J. A., Blackmore, K. M., Fan, J., Malone, K. E., John, E. M., Lynch, C. F., ... West, M. (2018). The association of mammographic density with risk of contralateral breast cancer and change in density with treatment in the WECARE study. *Breast Cancer Research*, 20(1), 1–10. <https://doi.org/10.1186/s13058-018-0948-4>
- Ko, K. L., Shin, I. S., You, J. Y., Jung, S. Y., Ro, J., & Lee, E. S. (2013). Adjuvant tamoxifen-induced mammographic breast density reduction as a predictor for recurrence in estrogen receptor-positive premenopausal breast cancer patients. *Breast Cancer Research and Treatment*, 142(3), 559–567. <https://doi.org/10.1007/s10549-013-2726-4>
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances In Neural Information Processing Systems*, 1–9. <https://doi.org/http://dx.doi.org/10.1016/j.protcy.2014.09.007>
- Kwok, S. M., Chandrasekhar, R., Attikiouzel, Y., & Rickard, M. T. (2006). Automatic pectoral muscle segmentation on mediolateral oblique view mammograms. *Recent Advances in Breast Imaging, Mammography, and Computer-Aided Diagnosis of Breast Cancer*, 23(9), 613–638. <https://doi.org/10.1117/3.651880.ch18>
- Lee, J., & Nishikawa, R. M. (2018a). Automated mammographic breast density estimation using

- a fully convolutional network: *Medical Physics*, 45(3), 1178–1190.  
<https://doi.org/10.1002/mp.12763>
- Lee, J., & Nishikawa, R. M. (2018b). Automated mammographic breast density estimation using a fully convolutional network: *Medical Physics*, 45(3), 1178–1190.  
<https://doi.org/10.1002/mp.12763>
- Lee, J., & Nishikawa, R. M. (2018c). *Automated mammographic breast density estimation using a fully convolutional network*. (October 2017), 1178–1190. <https://doi.org/10.1002/mp.12763>
- Lehman, C. D., Yala, A., Schuster, T., Dontchos, B., Bahl, M., Swanson, K., & Barzilay, R. (2018). Mammographic Breast Density Assessment Using Deep Learning: Clinical Implementation. *Radiology*, 180694. <https://doi.org/10.1148/radiol.2018180694>
- Lehman, C. D., Yala, A., Schuster, T., Dontchos, B., Bahl, M., Swanson, K., & Barzilay, R. (2019). Mammographic breast density assessment using deep learning: Clinical implementation. *Radiology*, 290(1), 52–58. <https://doi.org/10.1148/radiol.2018180694>
- Li, S., Wei, J., Chan, H., Helvie, M. A., Roubidoux, M. A., Lu, Y., & Zhou, C. (2018). *Computer-aided assessment of breast density : comparison of supervised deep learning and feature-based statistical learning Computer-aided assessment of breast density : comparison of supervised deep learning and feature-based statistical learning*.
- Li, S., Wei, J., Chan, H. P., Helvie, M. A., Roubidoux, M. A., Lu, Y., ... Samala, R. K. (2018). Computer-aided assessment of breast density: Comparison of supervised deep learning and feature-based statistical learning. *Physics in Medicine and Biology*, 63(2).  
<https://doi.org/10.1088/1361-6560/aa9f87>

- Linver, M. N. (2008). 4-19 Mammographic Density and the Risk and Detection of Breast Cancer. *Breast Diseases, 18*(4), 364–365. [https://doi.org/10.1016/S1043-321X\(07\)80400-0](https://doi.org/10.1016/S1043-321X(07)80400-0)
- Litjens, G., Kooi, T., Bejnordi, B. E., Arindra, A., Setio, A., Ciompi, F., ... Clara, I. S. (1998). *A Survey on Deep Learning in Medical Image Analysis*. (1995).
- Lokate, M., Stellato, R. K., Veldhuis, W. B., Peeters, P. H. M., & Van Gils, C. H. (2013). Age-related changes in mammographic density and breast cancer risk. *American Journal of Epidemiology, 178*(1), 101–109. <https://doi.org/10.1093/aje/kws446>
- Long, J., Shelhamer, E., & Darrell, T. (n.d.). *Fully Convolutional Networks for Semantic Segmentation*.
- Mac Parthaláin, N., Jensen, R., Shen, Q., & Zwiggelaar, R. (2010). Fuzzy-rough approaches for mammographic risk analysis. *Intelligent Data Analysis, 14*(2), 225–244. <https://doi.org/10.3233/IDA-2010-0418>
- Maskarinec, G., Pagano, I., Lurie, G., & Kolonel, L. N. (2006). A longitudinal investigation of mammographic density: The multiethnic cohort. *Cancer Epidemiology Biomarkers and Prevention, 15*(4), 732–739. <https://doi.org/10.1158/1055-9965.EPI-05-0798>
- Maskarinec, G., Woolcott, C. G., & Kolonel, L. N. (2010). *Cancer Outcome*. 351–354.
- Matsuyama, E., Takehara, M., & Tsai, D. (2020). *Using a Wavelet-Based and Fine-Tuned Convolutional Neural Network for Classification of Breast Density in Mammographic Images*. 17–29. <https://doi.org/10.4236/ojmi.2020.101002>
- McCormack, V. A. (2006). Breast Density and Parenchymal Patterns as Markers of Breast Cancer Risk: A Meta-analysis. *Cancer Epidemiology Biomarkers & Prevention, 15*(6), 1159–1169.

<https://doi.org/10.1158/1055-9965.EPI-06-0034>

McCormack, Valerie A., & Dos Santos Silva, I. (2006). Breast density and parenchymal patterns as markers of breast cancer risk: A meta-analysis. *Cancer Epidemiology Biomarkers and Prevention*, *15*(6), 1159–1169. <https://doi.org/10.1158/1055-9965.EPI-06-0034>

Mirjalili, S., & Lewis, A. (2016). The Whale Optimization Algorithm. *Advances in Engineering Software*, *95*, 51–67. <https://doi.org/10.1016/j.advengsoft.2016.01.008>

Mohamed, A. A., Berg, W. A., Peng, H., Luo, Y., Jankowitz, R. C., & Wu, S. (2018). A deep learning method for classifying mammographic breast density categories. *Medical Physics*, *45*(1), 314–321. <https://doi.org/10.1002/mp.12683>

Mohamed, A. A., Luo, Y., Peng, H., & Jankowitz, R. C. (2017). *Understanding Clinical Mammographic Breast Density Assessment : a Deep Understanding Clinical Mammographic Breast Density Assessment : a Deep Learning Perspective*. (September). <https://doi.org/10.1007/s10278-017-0022-2>

Mohamed, A. A., Luo, Y., Peng, H., & Jankowitz, R. C. (2018). *Understanding Clinical Mammographic Breast Density Assessment : a Deep Learning Perspective*. 387–392. <https://doi.org/10.1007/s10278-017-0022-2>

Moreira, I. C., Amaral, I., Domingues, I., Cardoso, A., Cardoso, M. J., & Cardoso, J. S. (2012). INbreast: Toward a Full-field Digital Mammographic Database. *Academic Radiology*, *19*(2), 236–248. <https://doi.org/10.1016/j.acra.2011.09.014>

Nagi, J., Abdul Kareem, S., Nagi, F., & Khaleel Ahmed, S. (2010). Automated breast profile segmentation for ROI detection using digital mammograms. *Proceedings of 2010 IEEE*

*EMBS Conference on Biomedical Engineering and Sciences, IECBES 2010*, (January), 87–92. <https://doi.org/10.1109/IECBES.2010.5742205>

Nickson, C., Arzhaeva, Y., Aitken, Z., Elgindy, T., Buckley, M., Li, M., ... Kavanagh, A. M. (2013). AutoDensity: An automated method to measure mammographic breast density that predicts breast cancer risk and screening outcomes. *Breast Cancer Research*, 15(5). <https://doi.org/10.1186/bcr3474>

Nyante, S. J., Sherman, M. E., Pfeiffer, R. M., Berrington De Gonzalez, A., Brinton, L. A., Aiello Bowles, E. J., ... Gierach, G. L. (2015). Prognostic significance of mammographic density change after initiation of tamoxifen for ER-positive breast cancer. *Journal of the National Cancer Institute*, 107(3), 5–11. <https://doi.org/10.1093/jnci/dju425>

Nyante, S. J., Sherman, M. E., Pfeiffer, R. M., Berrington De Gonzalez, A., Brinton, L. A., Aiello Bowles, E. J., ... Gierach, G. L. (2016). Longitudinal change in mammographic density among ER-positive breast cancer patients using tamoxifen. *Cancer Epidemiology Biomarkers and Prevention*, 25(1), 212–216. <https://doi.org/10.1158/1055-9965.EPI-15-0412>

Oliver, A., Lladó, X., Pérez, E., Pont, J., Denton, E. R. E., Freixenet, J., & Martí, J. (2010). A statistical approach for breast density segmentation. *Journal of Digital Imaging*, 23(5), 527–537. <https://doi.org/10.1007/s10278-009-9217-5>

Oliver, A., Tortajada, M., Lladó, X., Freixenet, J., Ganau, S., Tortajada, L., ... Martí, R. (2015). Breast Density Analysis Using an Automatic Density Segmentation Algorithm. *Journal of Digital Imaging*, 28(5), 604–612. <https://doi.org/10.1007/s10278-015-9777-5>

Orel, S. G., Kay, N., Reynolds, C., & Sullivan, D. C. (1999). BI-RADS Categorization As a Predictor of Malignancy. *Radiology*, 211(3), 845–850.

<https://doi.org/10.1148/radiology.211.3.r99jn31845>

Park, C. C., Rembert, J., Chew, K., Moore, D., & Kerlikowske, K. (2009). High Mammographic Breast Density Is Independent Predictor of Local But Not Distant Recurrence After Lumpectomy and Radiotherapy for Invasive Breast Cancer. *International Journal of Radiation Oncology Biology Physics*, 73(1), 75–79.

<https://doi.org/10.1016/j.ijrobp.2008.04.007>

Personal, I. (1996). Chandrasekhar , R . and Attikiouzel , Y . ( 1996 ) Gross segmentation of mammograms using a polynomial model . In : *Proceedings of the 18th Annual International Conference of the IEEE Engineering in Medicine and Biology Society . Bridging Disciplines for* . 1056–1058.

Petroudi, S., Constantinou, I., Pattichis, M., Tziakouri, C., Tziakouri, C., & Pattichis, C. (2015). Evaluation of spatial dependence matrices on multiscale instantaneous amplitude for mammogram classification. *IFMBE Proceedings*, 45, 156–159. [https://doi.org/10.1007/978-3-319-11128-5\\_39](https://doi.org/10.1007/978-3-319-11128-5_39)

Rampun, A., Scotney, B., Morrow, P., Wang, H., & Winder, J. (2018). Breast Density Classification Using Local Quinary Patterns with Various Neighbourhood Topologies. *Journal of Imaging*, 4(1), 14. <https://doi.org/10.3390/jimaging4010014>

Román, M., Sala, M., Baré, M., Posso, M., Vidal, C., Louro, J., ... Bargalló, X. (2019). Changes in mammographic density over time and the risk of breast cancer: An observational cohort study. *Breast*, 46, 108–115. <https://doi.org/10.1016/j.breast.2019.04.007>

Ronneberger, O., Fischer, P., & Brox, T. (n.d.). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. 1–8.



- Rouhi, R., Jafari, M., Kasaei, S., & Keshavarzian, P. (2015). Benign and malignant breast tumors classification based on region growing and CNN segmentation. *Expert Systems with Applications*, 42(3), 990–1002. <https://doi.org/10.1016/j.eswa.2014.09.020>
- Saffari, N., Rashwan, H. A., Abdel-Nasser, M., Singh, V. K., Arenas, M., Mangina, E., ... Puig, D. (2020). Fully automated breast density segmentation and classification using deep learning. *Diagnostics*, 10(11), 1–20. <https://doi.org/10.3390/diagnostics10110988>
- Saidin, N., Mat Sakim, H. A., Ngah, U. K., & Shuaib, I. L. (2013). Computer aided detection of breast density and mass, and visualization of other breast anatomical regions on mammograms using graph cuts. *Computational and Mathematical Methods in Medicine*, 2013. <https://doi.org/10.1155/2013/205384>
- Sanjay-Gopal, S., Chan, H. P., Wilson, T., Helvie, M., Petrick, N., & Sahiner, B. (1999). A regional registration technique for automated interval change analysis of breast lesions on mammograms. *Medical Physics*, 26(12), 2669–2679. <https://doi.org/10.1118/1.598806>
- Search, H., Journals, C., Contact, A., Iopscience, M., & Address, I. P. (n.d.). *The quantitative analysis of mammographic densities*. 1629. <https://doi.org/10.1118/1.598806>
- Seo, J. M., Ko, E. S., Han, B. K., Ko, E. Y., Shin, J. H., & Hahn, S. Y. (2013). Automated volumetric breast density estimation: A comparison with visual assessment. *Clinical Radiology*, 68(7), 690–695. <https://doi.org/10.1016/j.crad.2013.01.011>
- Setiawan, A. S., Elysia, Wesley, J., & Purnama, Y. (2015). Mammogram Classification using Law's Texture Energy Measure and Neural Networks. *Procedia Computer Science*, 59(Iccsci), 92–97. <https://doi.org/10.1016/j.procs.2015.07.341>

- Shelhamer, E., Long, J., & Darrell, T. (2017). Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4), 640–651. <https://doi.org/10.1109/TPAMI.2016.2572683>
- Simonyan, K., & Zisserman, A. (2014). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. 1–14. <https://doi.org/10.1016/j.infsof.2008.09.005>
- Sprague, B. L., Stout, N. K., Schechter, C., Van Ravesteyn, N. T., Cevik, M., Alagoz, O., ... Tosteson, A. N. A. (2015). Benefits, harms, and cost-effectiveness of supplemental ultrasonography screening for women with dense breasts. *Annals of Internal Medicine*, 162(3), 157–166. <https://doi.org/10.7326/M14-0692>
- Stewart, B. W., & Wild, C. P. (2015). *BOOK REVIEWS WORLD CANCER REPORT 2014*. XV(September), 2014–2016.
- Sun, L., Sc, M., Stone, J., Sc, M., & Fishell, E. (2007). *Mammographic Density and the Risk and Detection of Breast Cancer*. 227–236.
- Tai, S. C., Chen, Z. S., & Tsai, W. T. (2014). An automatic mass detection system in mammograms based on complex texture features. *IEEE Journal of Biomedical and Health Informatics*, 18(2), 618–627. <https://doi.org/10.1109/JBHI.2013.2279097>
- Titus-Ernstoff, L., Tosteson, A. N. A., Kasales, C., Weiss, J., Goodrich, M., Hatch, E. E., & Carney, P. A. (2006). Breast cancer risk factors in relation to breast density (United States). *Cancer Causes and Control*, 17(10), 1281–1290. <https://doi.org/10.1007/s10552-006-0071-1>
- Tzikopoulos, S. D., Mavroforakis, M. E., Georgiou, H. V., Dimitropoulos, N., & Theodoridis, S.

- (2011). A fully automated scheme for mammographic segmentation and classification based on breast density and asymmetry. *Computer Methods and Programs in Biomedicine*, *102*(1), 47–63. <https://doi.org/10.1016/j.cmpb.2010.11.016>
- Vachon, C. M., Ingle, J. N., Suman, V. J., Scott, C. G., Gottardt, H., Olson, J. E., & Goss, P. E. (2007). Pilot study of the impact of letrozole vs. placebo on breast density in women completing 5 years of tamoxifen. *Breast*, *16*(2), 204–210. <https://doi.org/10.1016/j.breast.2006.10.007>
- Vachon, Celine M., Pankratz, V. S., Scott, C. G., Maloney, S. D., Ghosh, K., Brandt, K. R., ... Sellers, T. A. (2007). Longitudinal trends in mammographic percent density and breast cancer risk. *Cancer Epidemiology Biomarkers and Prevention*, *16*(5), 921–928. <https://doi.org/10.1158/1055-9965.EPI-06-1047>
- Wang, A. T., Vachon, C. M., Brandt, K. R., & Ghosh, K. (2014). Breast density and breast cancer risk: A practical review. *Mayo Clinic Proceedings*, *89*(4), 548–557. <https://doi.org/10.1016/j.mayocp.2013.12.014>
- Wang, J., Yang, X., Cai, H., Tan, W., Jin, C., & Li, L. (2016). Discrimination of Breast Cancer with Microcalcifications on Mammography by Deep Learning. *Scientific Reports*, *6*(June), 1–9. <https://doi.org/10.1038/srep27327>
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, *13*(4), 600–612. <https://doi.org/10.1109/TIP.2003.819861>
- Whitman, G. J. (2010). Mammographic density and the risk of breast cancer recurrence after breast-conserving surgery. *Breast Diseases*, Vol. 21, pp. 226–227.

[https://doi.org/10.1016/S1043-321X\(10\)79560-6](https://doi.org/10.1016/S1043-321X(10)79560-6)

- Wolfe, J. N. (1976). Breast Patterns As Index of Breast. *AJR Am J Roentgenol*, 126, 1130–1139.
- Wu, S., Weinstein, S. P., Conant, E. F., & Kontos, D. (2013). Automated fibroglandular tissue segmentation and volumetric density estimation in breast MRI using an atlas-aided fuzzy C-means method. *Medical Physics*, 40(12), 1–12. <https://doi.org/10.1118/1.4829496>
- Youk, J. H., Gweon, H. M., Son, E. J., & Kim, J. A. (2016). Automated volumetric breast density measurements in the era of the BI-RADS fifth edition: A comparison with visual assessment. *American Journal of Roentgenology*, 206(5), 1056–1062. <https://doi.org/10.2214/AJR.15.15472>
- Zheng, Y., Keller, B. M., Ray, S., Wang, Y., Conant, E. F., Gee, J. C., & Kontos, D. (2015). Parenchymal texture analysis in digital mammography: A fully automated pipeline for breast cancer risk assessment. *Medical Physics*, 42(7), 4149–4160. <https://doi.org/10.1118/1.4921996>
- Zwiggelaar, R. (2010). Local greylevel appearance histogram based texture segmentation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6136 LNCS, 175–182. [https://doi.org/10.1007/978-3-642-13666-5\\_24](https://doi.org/10.1007/978-3-642-13666-5_24)



UNIVERSITAT  
ROVIRA I VIRGILI