

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS



**Enhanced Digital Breast Tomosynthesis diagnosis using 3D visualization and
automatic classification of lesions**

“ Documento Definitivo ”

Doutoramento em Engenharia Biomédica e Biofísica

Ana Margarida Pires de Almeida Mota

Tese orientada por:

Professor Doutor Pedro Almeida

Professor Doutor Matthew Clarkson

Professor Doutor Nuno Matela

Documento especialmente elaborado para a obtenção do grau de doutor

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS



**Enhanced Digital Breast Tomosynthesis diagnosis using 3D visualization and automatic
classification of lesions**

Doutoramento em Engenharia Biomédica e Biofísica

Ana Margarida Pires de Almeida Mota

Tese orientada por:

Professor Doutor Pedro Almeida

Professor Doutor Matthew Clarkson

Professor Doutor Nuno Matela

Júri:

Presidente:

- Doutor José Manuel de Nunes Vicente e Rebordão, Investigador Coordenador e Presidente do Departamento de Física da Faculdade de Ciências da Universidade de Lisboa

Vogais:

- Doutor João Manuel Ribeiro da Silva Tavares, Professor Catedrático da Faculdade de Engenharia da Universidade do Porto
- Doutor José Manuel Matos Ribeiro da Fonseca, Professor Associado com Agregação da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa
- Doutor Nuno Miguel de Pinto Lobo e Matela, Professor Auxiliar da Faculdade de Ciências da Universidade de Lisboa (orientador)
- Doutora Raquel Cruz da Conceição, Professora Auxiliar da Faculdade de Ciências da Universidade de Lisboa

Documento especialmente elaborado para a obtenção do grau de doutor

Entidade Financiadora: Universidade de Lisboa e Fundação para a Ciência e Tecnologia

This work was supported by Universidade de Lisboa (PhD grant) and Fundação para a Ciência e Tecnologia – Portugal (Grant No. SFRH/BD/135733/2018).



Nota Prévia

De acordo com o disposto no nº 5 do Artigo 3, Capítulo I, do Regulamento do Ciclo de Estudos Conducente ao Grau de Doutor da Faculdade de Ciências da Universidade de Lisboa, publicado no Diário da República, 2.^a série – N.º60 – 26 de Março de 2018, na elaboração desta tese foi efetuado um aproveitamento total dos resultados de trabalhos publicados ou aceites para publicação em revistas com comités de seleção de reconhecido mérito internacional.

The work presented in chapters 4 and 5 was also based on the following previous publications:

Ana M. Mota, Nuno Matela, Nuno Oliveira and Pedro Almeida. “Total variation minimization filter for DBT imaging”. *Medical Physics*. 2015 Jun; 42(6):2827-36. doi: 10.1118/1.4919680. PMID: 26127035.

Ana M. Mota, Nuno Matela, Nuno Oliveira and Pedro Almeida. “An iterative algorithm for Total Variation minimization in DBT imaging”. In *VipIMAGE 2015*, 2015; Tenerife, Spain: CRC Press. doi: <https://doi.org/10.1201/b19241>.

Ana M. Mota, Nuno Oliveira, Pedro Almeida and Nuno Matela. “3D Total Variation Minimization Filter for Breast Tomosynthesis Imaging”. In *Breast Imaging. IWDM 2016. Lecture Notes in Computer Science*, vol 9699. Springer, Cham. https://doi.org/10.1007/978-3-319-41546-8_63.

Acknowledgements

I would like to sincerely thank to:

Professor Pedro Almeida, for having always welcomed me at the IBEB with the greatest kindness and for guiding me with his knowledge and ambitious scientific goals.

Professor Nuno Matela, for all the patience, daily guidance and motivation he has given me not only in the PhD but also over the past few years.

Professor Matt Clarkson, for accepting my invitation to be one of my supervisors in this PhD thesis, without even knowing me. I really appreciate all the online meetings and the transmission of enormous scientific knowledge.

Professor Eduardo Ducla-Soares, for always being by my side when I needed it and for giving me precious advice, which I will always take with me in my academic and personal life.

Doctor Lurdes Orvalho, for all her medical expertise, patience and hours spent explaining to me the clinical point of view on what was really important in breast cancer detection.

Beatriz, for her generosity and friendship since the moment I joined the IBEB for the first time ten years ago.

Sofia and Nichal, my closest colleagues, with whom I shared every moment of this journey and who always supported me with their words and good spirits.

My Mother and Father, who are my anchor in all my life, professional and personal, and support me unconditionally.

Benjamin, my Husband, for all the love, understanding and resilience over these long years of PhD. Without him, it would certainly be difficult to achieve all the goals I had set myself.

And, of course, to my son **Rodrigo** and daughter **Júlia**, who were born during my PhD, making everything more challenging but even more valuable.

Abstract

Breast cancer represents the main cause of cancer-related deaths in women. Nonetheless, the mortality rate of this disease has been decreasing over the last three decades, largely due to the screening programs for early detection. For many years, both screening and clinical diagnosis were mostly done through Digital Mammography (DM). Approved in 2011, Digital Breast Tomosynthesis (DBT) is similar to DM but it allows a 3D reconstruction of the breast tissue, which helps the diagnosis by reducing the tissue overlap. Currently, DBT is firmly established and is approved as a stand-alone modality to replace DM.

The main objective of this thesis is to develop computational tools to improve the visualization and interpretation of DBT data.

Several methods for an enhanced visualization of DBT data through volume rendering were studied and developed. Firstly, important rendering parameters were considered. A new approach for automatic generation of transfer functions was implemented and two other parameters that highly affect the quality of volume rendered images were explored: voxel size in z direction and sampling distance. Next, new image processing methods that improve the rendering quality by considering the noise regularization and the reduction of out-of-plane artifacts were developed.

The interpretation of DBT data with automatic detection of lesions was approached through artificial intelligence methods. Several deep learning Convolutional Neural Networks (CNNs) were implemented and trained to classify a complete DBT image for the presence or absence of microcalcification clusters (MCs). Then, a faster R-CNN (region-based CNN) was trained to detect and accurately locate the MCs in the DBT images. The detected MCs were rendered with the developed 3D rendering software, which provided an enhanced visualization of the volume of interest. The combination of volume visualization with lesion detection may, in the future, improve both diagnostic accuracy and also reduce analysis time.

This thesis promotes the development of new computational imaging methods to increase the diagnostic value of DBT, with the aim of assisting radiologists in their task of analyzing DBT volumes and diagnosing breast cancer.

Keywords: Breast cancer; Digital Breast Tomosynthesis; 3D volume rendering; artificial intelligence; microcalcifications.

Resumo

O cancro da mama é um dos exemplos onde os avanços da tecnologia e terapêutica têm feito uma enorme diferença na redução da mortalidade. Apesar de ser o tipo de cancro com maior incidência e de ainda representar a principal causa de morte por cancro nas mulheres, esta taxa de mortalidade tem vindo a diminuir ao longo das últimas três décadas. Este declínio está diretamente relacionado com a implementação e sucesso dos programas de rastreio para uma deteção precoce de cancro da mama (nomeadamente nos países desenvolvidos). É sabido que, quanto mais cedo uma lesão maligna for detetada, menor é a probabilidade desta se desenvolver num cancro mais agressivo e metastizado, aumentando assim a sobrevida das mulheres diagnosticadas.

Até há cerca de dez anos, o principal meio de diagnóstico de cancro da mama era a mamografia, quer em rastreio, quer na prática clínica diária. Num exame de mamografia, a mama é comprimida entre duas placas e são adquiridas duas projeções de raios-x de baixa dose. Esta técnica de imagiologia apresenta uma enorme desvantagem: a sua aquisição bidimensional (2D) resulta numa imagem da sobreposição dos tecidos mamários, sem qualquer informação de profundidade dos mesmos. Este facto levanta essencialmente dois problemas: (1) alguns falsos negativos, onde lesões são escondidas por tecidos mais densos circundantes ou (2) elevado número de falsos positivos, nos quais a sobreposição de tecidos imita uma lesão. Ambas as situações são bastante prejudiciais para as pacientes uma vez que, no primeiro caso, há uma lesão maligna que pode não ser detetada atempadamente, podendo originar posteriormente um cancro mais agressivo. Na segunda situação, a paciente tem de fazer mais exames complementares e biópsias, resultando numa enorme angústia e ansiedade, bem como custos financeiros adicionais que seriam completamente desnecessários.

Desta forma, em 2011 foi aprovada uma nova técnica de imagiologia com base na mamografia mas que permite adquirir várias projeções da mama e, assim, reconstruir um objeto tridimensional (3D) com alguma noção de profundidade: a Tomossíntese Mamária (TM). Com a TM, é possível analisar o tecido mamário adquirido ao longo de vários cortes e, assim, “separar” melhor a informação obtida em profundidade. Este exame é realizado atualmente com aproximadamente a mesma dose de radiação para o paciente que um exame convencional de mamografia. Apesar de ser uma técnica relativamente recente, já consolidou o seu papel na imagiologia do cancro da mama. Inicialmente começou por ser realizada em conjunto com a mamografia, quer para ser possível uma comparação com mamografias anteriores, quer porque ainda decorriam

estudos que comprovassem a verdadeira mais-valia de TM na substituição da mamografia. A utilização de TM+mamografia demonstrou um aumento nas taxas de detecção do cancro da mama e uma redução significativa nas taxas de rechamada, particularmente em mulheres com mamas densas. Atualmente, com a possibilidade de gerar uma mamografia sintética através dos dados de TM (ultrapassando assim o problema de comparar mamografias da mesma mama ao longo do tempo), a TM está aprovada como uma modalidade de imagiologia autónoma que pode substituir a mamografia convencional 2D. No entanto, como qualquer técnica recente, todos os aspetos da sua utilização têm ainda lacunas que devem ser exploradas e corrigidas, passando pelo *hardware*, aquisição, reconstrução, visualização e interpretação de dados.

O trabalho desenvolvido nesta tese teve como objetivo principal o desenvolvimento de ferramentas computacionais para a visualização e interpretação de dados da TM. Todo o *software* de visualização foi escrito em C++ utilizando a biblioteca *open-source* VTK (do inglês, Visualization Toolkit). Os algoritmos de processamento e de inteligência artificial foram implementados em MATLAB.

Atualmente, a visualização dos dados é feita através de um modo corte a corte ou num *loop* contínuo onde os radiologistas exploram e avaliam cada exame. Tal procedimento permite uma verificação de estruturas anatómicas apenas em duas dimensões, enquanto a terceira é reconstruída mentalmente pelo radiologista olhando para cortes adjacentes. Além disso, há lesões, como aglomerados de microcalcificações, que podem estar dispersos por vários cortes em profundidade, dificultando a análise da sua distribuição espacial e, assim, a sua interpretação e diagnóstico. Por outro lado, como um exame de TM tem aproximadamente 60 imagens de corte para analisar, em comparação com apenas uma da mamografia convencional, avaliar cada exame de TM leva a uma inspeção mais demorada. Assim, uma forma de visualização de dados de TM diferente poderá desempenhar um papel complementar importante, como é feito nos casos da Tomografia Computorizada ou da Imagiologia por Ressonância Magnética.

A visualização por renderização volumétrica consiste na criação de uma cena 3D através de imagens computacionais realistas, produzindo uma verdadeira perceção de profundidade. Através da observação do objeto segundo vários ângulos (e não apenas num plano da imagem, como acontece com a visualização corte a corte), é possível fazer uma inspeção global imediata e ter uma compreensão dos dados subjacentes de uma só vez. Neste trabalho foram desenvolvidas técnicas e ferramentas necessárias para uma visualização adequada dos dados de TM através de renderização volumétrica.

As funções de transferência (que atribuem opacidade e cor a certos elementos, como valores de intensidade e/ou de magnitude de gradiente) são um parâmetro crítico na renderização, afetando diretamente o que se vê e de que forma. Assim, encontrar funções de transferência adequadas é um dos maiores desafios da renderização volumétrica. Por esse motivo, foi desenvolvida metodologia que permitisse gerar estas funções de forma automática, resultando em renderizações úteis e apropriadas dos dados de TM. A abordagem implementada permitiu a produção automática de várias funções de opacidade/cor Vs. intensidade escalar e magnitude de gradiente. Adicionalmente, este trabalho explorou outros dois parâmetros que também afetam a qualidade da imagem renderizada: dimensão dos voxels na direção z e a distância de amostragem. Como os dados reconstruídos de TM têm voxels anisotrópicos, ao efectuar uma rotação em 3D em torno do objeto, observa-se distorção e perda de qualidade. Assim, para estudar o efeito de voxels isotrópicos neste tipo de visualização, o tamanho de voxel foi alterado de $0,085 \times 0,085 \times 1,0 \text{ mm}^3$ para $0,085 \times 0,085 \times 0,085 \text{ mm}^3$ usando várias funções de interpolação da biblioteca VTK. Por outro lado, vários valores de distância de amostragem foram introduzidos no *software* de visualização desenvolvido e a sua influência na qualidade da imagem renderizada foi avaliada.

Além dos parâmetros referidos, também foi necessário desenvolver técnicas de processamento de imagem que melhorassem o aspeto da renderização em dois pontos: (1) diminuição de ruído e (2) redução do artefacto “out-of-plane”. Para o primeiro ponto foram implementados algoritmos 3D de minimização da variação total dos dados (do inglês, Total Variation – TV), uma vez que estes têm demonstrado excelentes resultados na diminuição do ruído enquanto preservam as bordas dos objetos. Para o segundo ponto foi desenvolvida uma nova abordagem em TM para redução de artefactos “out-of-plane” usando desconvolução cega (do inglês, blind deconvolution) e regularização do ruído através de minimização da variação total. Com os parâmetros de renderização apropriados e as duas metodologias de processamento desenvolvidas especificamente para melhorar as imagens renderizadas, este trabalho alcançou uma melhoria significativa da qualidade da renderização volumétrica de dados de TM.

Como mencionado, além da visualização, este trabalho focou-se também na interpretação dos dados de TM através de classificação e detecção automática de lesões. Até recentemente, apesar dos esforços e desenvolvimentos dos sistemas convencionais de Detecção Assistida por Computador (do inglês, Computer-Aided Detection – CAD) para auxiliar na interpretação de exames de TM, as suas elevadas taxas de falsos positivos e baixa especificidade, fizeram com que não

alcançassem um nível de desempenho que pudesse ser traduzido numa verdadeira melhoria na detecção do cancro da mama. Nos últimos anos, o aumento do poder computacional e o crescimento das bases de dados têm permitido o desenvolvimento de algoritmos de inteligência artificial de aprendizagem profunda compostos por redes neuronais convolucionais (do inglês, Convolutional Neural Networks - CNNs) de várias camadas. Estes algoritmos têm surgido como uma potencial solução para a detecção/interpretação automatizada do cancro da mama. Desta forma, numa primeira fase, foram implementadas e treinadas várias CNNs e o seu desempenho foi avaliado no que respeita à sua capacidade para classificar uma imagem completa de TM quanto à presença ou ausência de aglomerados de microcalcificações (sem qualquer identificação prévia de regiões candidatas). Numa segunda fase, uma *faster* R-CNN (do inglês, faster region-based CNN) foi treinada para detetar e localizar os aglomerados de microcalcificações nas imagens de TM. Depois de detetados, a informação acerca da localização destes aglomerados é fornecida ao *software* de renderização 3D desenvolvido previamente e, ao invés de toda a mama ser renderizada, é possível visualizar apenas o volume de interesse contendo as microcalcificações detetadas. Esta parte do trabalho focou-se neste tipo de lesões uma vez que os aglomerados de microcalcificações são um dos biomarcadores mais importantes para o cancro da mama, especialmente em casos de lesões não palpáveis. Além disso, a sua reduzida dimensão requer bastante resolução espacial e, portanto, maior poder computacional, fazendo com que a maioria dos trabalhos desenvolvidos com inteligência artificial em TM até à data fosse acerca de lesões de tecidos moles, como massas ou distorções arquiteturais.

Este trabalho promoveu assim a investigação e desenvolvimento de novos métodos de imagem computacional para aumentar o valor de diagnóstico de TM, com o objetivo de auxiliar os radiologistas na sua tarefa de analisar volumes de TM e diagnosticar cancro de mama.

Palavras-chave: Cancro da mama; Tomossíntese mamária; renderização volumétrica; inteligência artificial; microcalcificações.

Table of Contents

Acknowledgements	XI
Abstract	XIII
Resumo.....	XV
List of Acronyms.....	XXI
List of Figures	XXIII
List of Tables.....	XXVII
Introduction	1
1.1. Breast cancer	2
1.2. Digital Breast Tomosynthesis.....	4
1.3. Visualization of DBT.....	6
1.4. Automatic classification and detection of lesions.....	9
1.5. Thesis outline	12
References	15
Calculation of transfer functions for volume rendering of breast tomosynthesis imaging	21
2.1. Introduction	23
2.2. Methods.....	24
2.3. Results	25
2.4. Discussion and conclusion.....	30
References	31
Optimization of Breast Tomosynthesis Visualization through 3D Volume Rendering.....	33
3.1. Introduction	35
3.2. Materials and methods.....	37
3.3. Results	41
3.4. Discussion	51
3.5. Conclusions	54
References	55
Impact of total variation minimization in volume rendering visualization of breast tomosynthesis data	57
4.1. Introduction	60
4.2. Methods.....	62
4.3. Results	67
4.4. Discussion	73
4.5. Conclusions	76
References	77
An enhanced visualization of DBT imaging using blind deconvolution and total variation minimization regularization	81

5.1. Introduction	83
5.2. Materials and methods.....	85
5.3. Results	89
5.4. Discussion	94
5.5. Conclusion.....	97
References	98
Automatic classification of simulated breast tomosynthesis whole images for the presence of microcalcification clusters using deep learning CNNs	101
6.1. Introduction	104
6.2. Materials and methods.....	108
6.3. Results	113
6.4. Discussion	117
6.5. Conclusions	121
References	122
Detection of microcalcifications in Digital Breast Tomosynthesis using Faster R-CNN and 3D volume rendering.....	129
6.1. Introduction	131
7.2. Materials and methods.....	132
7.3. Results	136
7.4. Discussion and conclusions	141
References	144
Final considerations.....	147

List of Acronyms

2D	Two dimensional
3D	Three dimensional
AI	Artificial Intelligence
ASF	Artifact spread function
AUC	Area under the curve
BF	Blur factor
BG	Background
BI-RADS	Breast Imaging Reporting and Data System
CADe	Computer Aided Detection
CADx	Computer Aided Diagnosis
CNN	Convolutional Neural Network
CNR	Contrast to noise ratio
CT	Computed Tomography
DBT	Digital Breast Tomosynthesis
DM	Digital Mammography
FBP	Filtered back projection
FDA	Food and Drug Administration
FWHM	Full width at half maximum
MC	Microcalcification cluster
MIP	Maximum Intensity Projection
PSF	Point spread function
ROI	Region of interest
SM	Synthetic Mammography
SNR	Signal to noise ratio
SSIM	Structural similarity index
TF	Transfer function
TV	Total variation
VR	Volume rendering
VTK	Visualization toolkit
WHW	Window half-width
λ	Lagrange multiplier

List of Figures

Figure 1.1. Examples of typically benign breast masses and suspected malignant masses.	3
Figure 1.2. Examples of typically benign microcalcifications (up) and suspected malignant microcalcifications (down).....	3
Figure 1.3. Examples of typically benign architectural distortions (up) and suspected malignant distortions (down).	3
Figure 1.4. 2D DM Acquisition schematic.....	4
Figure 1.5. 3D DBT Acquisition schematic.	5
Figure 1.6. Example of a cancer detected by DBT alone and missed with DM.....	6
Figure 1.7. Example showing that structures at different levels in the breast can summate to create suspicious region on DM that may be identified as negative or superimposed tissue on DBT images. Left: DM with a suspicious area (arrow); right: three slices of DBT with no suspicious area.....	6
Figure 1.8. Illustrative scheme of MIP (a) and compositing (b) technique.....	8
Figure 1.9. Examples of magnetic resonance rendered images obtained with (a) MIP and (b) the compositing technique.....	8
Figure 1.10. Illustration of the sequential arrangement of layers in a very simple CNN.....	11
Figure 2.1. Left: 1D linear ramp transfer function used to map voxel intensity values to opacity and color. Right: Zoom in on the region 1 and region 2 whose intensities correspond to a_1 and b_1 , respectively.....	24
Figure 2.2. Linear regressions obtained between a_1 and b_1 and (a) mean, (b) maximum and (c) skewness of histogram intensity values, calculated for each clinical case. Coefficients of determination (R^2) and equations for the estimation of a_1 and b_1 based on each statistic are also shown.	26
Figure 2.3. 2D displays of VR visualizations at 0° obtained with each combination of intensity-based TFs by varying the color and opacity within the values of a_1 , a_2 , b_1 and b_2	27
Figure 2.4. 2D displays of VR visualizations at 0° obtained with the ten gradient-based TFs: linear ramp (a), square shaped centered on each peak (b,c,d), ramp-shaped around each peak (e,f,g) and triangular shaped centered on each peak (h,i,j). Opacity and color values were kept at 1 for all scalar intensities.	28
Figure 2.5. 2D displays of VR visualizations at 0° considering intensity and gradient magnitude functions selected. 1st, 2nd and 3rd rows were obtained using the gradient-based functions: linear ramp and triangular-shaped centered on peak 2 and 3, respectively. 1st, 2nd and 3rd columns were obtained using the scalar intensity-based functions: 1, 2 and 6, respectively.	29
Figure 3.1. (a) Acrylic phantom simulating breast tissue and high density lesions (aluminum disks of different diameters and 1 mm thickness). (b) Scheme of the disks in the first column (top to bottom): 5.0 mm, 3.0 mm, 1.0 mm, 0.5 mm, 2.0 mm, and 4.0 mm, respectively.....	37
Figure 3.2. Illustrative scheme of visualization at 0° and 90°	38
Figure 3.3. Computation time required for rendering the original data (z : 1.0 mm) and data after rescaling with linear, cubic, nearest interpolators and Lanczos, Kaiser, Cosine, Hann, Hamming, Blackman, and Nuttall window functions (with sampling distance 1.0 mm).....	41
Figure 3.4. Profile of the 5.0 mm disk obtained at 90° for the original data (z : 1.0 mm) and after rescaling with the linear, cubic, and nearest interpolators. Zoom-in of a range with large intensity variation. ...	42
Figure 3.5. Profiles of the 5.0 mm disk obtained at 90° for the original data (z : 1.0 mm) and after rescaling with Lanczos, Kaiser, Cosine, Hann, Hamming, Blackman, and Nuttall window functions with WHW values of 1, 3, 5, 8, 13, and 16 ((a) to (f), respectively). Zoom-in of a range with large intensity variation.....	43

Figure 3.6. Results obtained with Lanczos, Kaiser, Cosine, Hann, Hamming, Blackman, and Nuttall window functions for the profile at 90°, by changing WHW values (from 1 to 16). (a) Smoothness values as a function of total time and (b) FWHM of the 5.0 mm disk at 90°.....	44
Figure 3.7. Profiles of the 5-mm disk obtained at 90° for the original data (z: 1.0 mm) and after rescaling with Lanczos, Kaiser, Cosine, Hann, Hamming, Blackman, and Nuttall window functions with BF(z) values of 1.0, 1.5, 2.0, 2.5, 3.0, and 4.0 ((a)–(f)). Zoom-in of a range with large intensity variation...	45
Figure 3.8. Results obtained with Lanczos, Kaiser, Cosine, Hann, Hamming, Blackman, and Nuttall window functions for the profile at 90°, by changing BF(z) values (from 1.0 to 4.0). (a) Smoothness values as a function of total time and (b) FWHM of the 5.0 mm disk at 90°.....	46
Figure 3.9. Computation time required for rendering the original data (black) and data after rescaling (gray) taking into account the different sampling distance values. For each sampling distance, each gray value was obtained by averaging the rendering times recorded for each interpolator considered here (cubic, Hamming with WHW = 5, Hamming with BF (z) = 2 and Hamming with WHW = 5 and BF (z) = 2).	47
Figure 3.10. Profiles of the 5.0 mm disk obtained at 90°, for some values of sampling distance tested (0.010, 0.025, 0.050, 0.100, 0.170, 0.195, and 1.0 mm), with the original data (a) and after rescaling with cubic (b), Hamming with WHW = 5 (c), Hamming with BF (z) = 2 (d), and Hamming with WHW = 5 and BF (z) = 2 (e).....	48
Figure 3.11. Smoothness (a), CNR (b) and FWHM (c) plotted as a function of sampling distance values for original data and data after interpolation with cubic, Hamming with WHW = 5, Hamming with BF (z) = 2 and Hamming with WHW = 5 and BF (z) = 2. Results obtained for rendered images at 90°...	49
Figure 3.12. Volume rendering images at 0° and 90° for the 5-mm disk obtained for the original data with default visualization (top row) and interpolated data with Hamming window with BF (z) = 2 and sampling distance 0.025 mm (bottom row).....	50
Figure 3.13. 2D displays of composite volume rendering visualization obtained at 0° and 90° (a and b, respectively) for original data with default sampling distance (1.0 mm) (a1 and b1) and interpolated data with sampling distance 0.025 mm (a2 and b2).....	51
Figure 4.1. Acrylic phantom simulating breast tissue and lesions of high attenuation (aluminum disks of different diameters and 1 mm thickness). Diameter of the first column disks (top to bottom): 5 mm, 3 mm, 1 mm, 0.5 mm, 2 mm and 4 mm, respectively.	62
Figure 4.2. Illustrative scheme of the application of each filter in slice number two. (a) 2D TV minimization filter applied to all slices, one slice at a time. (b) 3D TV minimization filter.....	65
Figure 4.3. 3D TV values for unfiltered and filtered phantom data, with minimization of 2D TV in all slices, plotted as a function of λ . Right: zoom in on λ range where 3D TV is minimum.....	68
Figure 4.4. 3D TV values for unfiltered and filtered phantom data, with minimization of 3D TV, plotted as a function of λ . Right: zoom in on λ range where 3D TV is minimum.	68
Figure 4.5. Values of CNR at 90° obtained with 2D TV minimization algorithm applied to all slices (a) and 3D TV minimization algorithm (b) as a function of λ	69
Figure 4.6. Volume rendering images at 0° for 5 mm disk (top row) and 0.5 mm disk (bottom row) obtained for the unfiltered (a1 and a2) and filtered data with 2D TV minimization filter applied to all slices – $\lambda=197$ (b1 and b2) and 3D TV minimization filter – $\lambda=195$ (c1 and c2).	70
Figure 4.7. Volume rendering images at 90° for 5 mm disk (top row) and 0.5 mm disk (bottom row) obtained for the unfiltered (a1 and a2) and filtered data with 2D TV minimization filter applied to all slices – $\lambda=197$ (b1 and b2) and 3D TV minimization filter – $\lambda=195$ (c1 and c2).	71
Figure 4.8. 3D TV values for unfiltered and filtered clinical data, with minimization of 2D TV in all slices, plotted as a function of λ . Right: zoom in on λ range where 3D TV is minimum.....	71

Figure 4.9. 3D TV values for unfiltered and filtered clinical data, with minimization of 3D TV, plotted as a function of λ . Right: zoom in on λ range where 3D TV is minimum.	72
Figure 4.10. Volume rendering at 0° (a, b and c) and 90° (d, e and f) of clinical data. (a1, b1, c1, d1, e1 and f1) represent the zoom in of the biggest calcification. Unfiltered data: a (a1) and d (d1); filtered data with 2D TV minimization: b (b1) and e (e1); and data obtained with 3D TV minimization filter: c (c1) and f (f1).	73
Figure 5.1. (a) Scheme of the acrylic phantom used. (b) Image of the phantom simulating breast tissue and lesions of high attenuation (aluminum disks of different diameters and 1 mm thickness). (b) Scheme of the disks in the first column (top to bottom): 5.0 mm, 3.0 mm, 1.0 mm, 0.5 mm, 2.0 mm and 4.0 mm, respectively.	85
Figure 5.2. Pseudocode of the four methods implemented in order to study the application of blind deconvolution in the reduction of out-of-plane artifacts in DBT data.	87
Figure 5.3. SSIM values obtained for deblurred data with each iterative blind deconvolution method, as a function of iteration number.	89
Figure 5.4. ASF curves obtained for each deconvolution method for the 0.5 mm (a), 1.0 mm (b), 2.0 mm (c), 3.0 mm (d), 4.0 mm (e) and 5.0 mm (f) disk.	90
Figure 5.5. FWHM _{ASF} values of a Gaussian curve fitted to the ASF of each disk (0.5 mm, 1.0 mm, 2.0 mm, 3.0 mm, 4.0 mm and 5.0 mm), for each method and original data.	91
Figure 5.6. FWHM _{90°} values of a Gaussian curve fitted to the profile of each disk at 90° (0.5 mm, 1.0 mm, 2.0 mm, 3.0 mm, 4.0 mm and 5.0 mm), for each method and original data.	91
Figure 5.7. Noise at 0° (a) and 90° (b) obtained for phantom original and deblurred data with each method plotted as a function of disk diameter.	92
Figure 5.8. SNR at 0° (a) and 90° (b) obtained for phantom original and deblurred data with each method plotted as a function of disk diameter.	92
Figure 5.9. 2D displays of composite volume rendering visualization obtained at 0° for blurred (a) and deblurred data with method 1 (b), method 2 (c), method 3 (d) and method 4 (e).	93
Figure 5.10. 2D displays of composite volume rendering visualization obtained at 90° for blurred (a) and deblurred data with method 1 (b), method 2 (c), method 3 (d) and method 4 (e).	93
Figure 5.11. 2D displays of composite volume rendering visualization obtained at 0 and 90 degrees (a and b, respectively) for blurred (a1 and b1) and deblurred data with method 4 (a2 and b2).	94
Figure 6.1. The six pre-processing methodologies implemented in order to reduce noise and amplify the visibility of the MCs (BG: background, normData: Data normalized between 0 and 1).	109
Figure 6.2. Illustration of CNN-a that resulted from the modifications made (bold) to the Alexnet architecture. <i>Conv</i> and <i>GroupConv</i> : convolutional and grouped convolutional layers, respectively; <i>pool</i> : max pooling layers; <i>fc</i> : fully connected layer; <i>relu</i> : rectified linear unit layer; <i>norm</i> : batch normalization layer; <i>drop</i> : dropout layer.	111
Figure 6.3. Summary of the methodological pipeline followed in this work.	112
Figure 6.4. a) Data with contaminated BG; b) First binary image; c) Filled binary image; d) Largest object extracted from binary image; e) Result from region growing; f) Final image with BG corrected after binary mask from e) applied to a).	113
Figure 6.5. a) Original data without pre-processing; b) Pre-processing 1 (minimization of TV); c) Pre-processing 2 (CLAHE); d) Pre-processing 3 (minTV + CLAHE); e) Pre-processing 4 (CLAHE + minTV); f) Pre-processing 5 (dataNorm2); g) Pre-processing 6 (dataNorm2 + minTV).	114
Figure 6.6. Comparisons of ROC curves for the CNNs and training data with the best AUC values.	115

Figure 6.7. Values of sensitivity, specificity and accuracy obtained with the architectures trained with pre-processed data that achieved the best mean AUC.....	116
Figure 6.8. AUC values obtained with test datasets composed by the four different breast densities separately (* $p < 0.05$: there is a significant difference between groups).....	116
Figure 6.9. Some examples of MCs in the DBT data used. (a) True positive (case classified correctly as positive by all CNNs, even in the original image); (b) False negative (case incorrectly classified as negative by all CNNs, even varying the pre-processing); (c) original case classified as negative and that was only detected by Googlenet when pre-processed with method 3 (d); (e) original case classified as negative and that was only detected by CNN-a when pre-processed with method 6 (f).....	119
Figure 7.1. Pipeline followed for connection between the output of the trained faster R-CNN and the volume rendering visualization of the detected object.	135
Figure 7.2. The FROC curve for the test dataset.	136
Figure 7.3. Example of four detection outputs obtained with a threshold of 0.9. Green: Ground truth BB; Yellow: predicted BB (without score: FPs, with score: true positives). The predicted results are visualized with 2D slice-by-slice represented through xy and xz planes and 3D VR with five different angles (0°, 22.5°, 45°, 67.5° and 90°).	138
Figure 7.4. Example of four missed detections (false negatives) and four incorrect detections (FPs). The BB are visualized through 2D slice-by-slice in xy and 3D VR at 0°.	139
Figure 7.5. Example of four detection outputs. Green: Ground truth BB; yellow: predicted BB. The predicted results are visualized with 2D MIP slice-by-slice represented through xy and xz planes and 3D VR with two different angles (0° and 90°). Each 3D VR was obtained using two different transfer functions, allowing different levels of MC segmentations.....	140

List of Tables

Table 1.1. Some characteristics of FDA approved DBT systems.....	5
Table 3.1. Summary of the functions and parameters, available in VTK, analyzed during the interpolation process.....	40
Table 3.2. FWHM and smoothness values measured in interpolated data at 90°. The total time (interpolation plus rendering time) is also presented.	42
Table 3.3. Summary of the results obtained for phantom at 0° and 90°, with default visualization options (voxels with $0.085 \times 0.085 \times 1.0$ mm ³ and sampling distance 1.0 mm) and the options selected in our study (voxels with $0.085 \times 0.085 \times 0.085$ mm ³ after interpolation with Hamming window function and $BF(z) = 2$ and sampling distance 0.025 mm).	50
Table 4.1. Summary of the differences between the previous studies using the mentioned TV minimization algorithms [42, 44] and the studies developed in this work.	66
Table 4.2. Summary of the results obtained for the 3D TV values of unfiltered and filtered phantom data with the two applications (minimization of 2D TV in all slices and minimization of 3D TV). It is also presented the variation in percentage between the unfiltered and filtered values.....	68
Table 4.3. Results obtained for CNR, FWHM _{0.5mm} and FWHM _{5.0mm} values of unfiltered and filtered phantom data at 90°. Each filtered data was obtained using λ corresponding to the maximum CNR at 90° ($\lambda=197$ for the 2D TV minimization in all slices and $\lambda=195$ for the 3D TV minimization). It is also presented the variation in percentage between the unfiltered and filtered values.	69
Table 4.4. Results obtained for CNR, FWHM _{0.5mm} and FWHM _{5.0mm} values of unfiltered and filtered phantom data at 0°. Each filtered data was obtained using λ corresponding to the maximum CNR at 90° ($\lambda=197$ for the 2D TV minimization in all slices and $\lambda=195$ for the 3D TV minimization). It is also presented the variation in percentage between the unfiltered and filtered values.	69
Table 4.5. Summary of the results obtained for the 3D TV values of unfiltered and filtered clinical data with the two applications (minimization of 2D TV in all slices and minimization of 3D TV). It is also presented the variation in percentage between the unfiltered and filtered values.	72
Table 4.6. Results obtained for CNR and FWHM _{Calc} values of unfiltered and filtered clinical data at 0° and 90°. Each filtered data was obtained using λ corresponding to the minimum 3D TV ($\lambda=135$ for the 2D TV minimization in all slices and $\lambda=145$ for the 3D TV minimization). It is also presented the variation in percentage between the unfiltered and filtered values.	72
Table 5.1. Variation, in percentage, between the FWHM _{ASF} values achieved for the original data and after deconvolution with each method. Taking into account the different disks, the average of the obtained variations (reductions) with each method is presented in the last row.	91
Table 5.2. Variation, in percentage, between the FWHM _{90°} values achieved for the original data and after deconvolution with each method. Taking into account the different disks, the average of the obtained variations with each method is presented in the last row.	91
Table 6.1. Summary of deep learning DBT studies (ROI: Region of interest, AUC: Area under the curve, pAUC: partial AUC).	105
Table 6.2. Detailed summary of the VICTRE data selected for this study.....	108
Table 6.3. Performance results of CNNs trained with original data and with data resulting from the pre-processing methodologies, in terms of mean AUC.	114
Table 6.4. Levels of significance (p-values) obtained from the statistical analysis of the difference between the best mean AUCs found.	115
Table 6.5. Training times, in hours, needed for each CNN after 3-fold cross validation and mean inference time (in seconds) needed to classify each image.	117

Table 7.1. Parameters used to design the Faster R-CNN.....	134
Table 7.2. Options used to train the Faster R-CNN.....	134
Table 7.3. The sensitivity values for less than 8 FP/image.....	136

Introduction

1

1.1. Breast cancer

Along with cardiovascular disease, cancer is a major public health concern today and the number of new cancer cases is expected to continue to increase over the next two decades. In 2020, an estimated 19.3 million new cancer cases and almost 10.0 million cancer deaths occurred worldwide. Cancer represents one of the four leading causes of premature death under the age of 70 years [1]. Currently, among all cancers and both sexes, breast cancer is the one with the highest incidence rate (followed by lung and prostate cancer), representing the fourth highest mortality rate [1, 2]. Although breast cancer mortality has been decreasing by about 40% from its peak in 1989, it still represents the biggest cause of cancer mortality among women [2]. The simultaneous increase in incidence and reduction in mortality are directly related to the increase in early stage-detection of this type of pathology. This early detection is possible due to the implemented screening programs, which are crucial in identifying early stage breast cancers that would, otherwise, be detected later when they could be more invasive and have metastasized. This initial intervention is extremely important to obtain a better outcome of treatments. It is estimated that women who chose to participate in an organized breast cancer screening program have 60% lower risk of dying from breast cancer within 10 years after diagnosis [3]. These screening programs are implemented all over the world, predominantly in developed countries. Asymptomatic women aged mainly between 50 and 69 years should have an x-ray imaging exam every two years to detect possible unknown lesions [4].

The main lesions and abnormalities related to breast cancer that are searched by the radiologists in these screening exams are: masses, microcalcifications and architectural distortions [5]. A mass results from an abnormal aggregation of breast tissue cells and is defined as a lesion seen in at least two different projections (if a suspicious mass is seen in only a single projection, it should be called an “asymmetry”). Masses are often difficult to detect, sometimes becoming imperceptible, because their features are similar to those of normal breast tissue (Figure 1.1). Microcalcifications are very small calcium deposits in breast tissue that appear as bright spots on images acquired with x-rays. Although most are considered benign, if they appear in the form of a cluster with a specific pattern, microcalcifications are more likely a sign of cancer [6]. Indeed, microcalcifications are often the only indication of the presence of breast cancer, essentially in non-palpable lesions (Figure 1.2) [7]. Architectural distortions of the parenchyma occurs when the normal architecture of the breast tissue is distorted with no visible mass. Although its frequency is small compared to masses or microcalcifications, it is also more

difficult to diagnose as it can be very subtle and be presented in a variety of forms (Figure 1.3). Proof of this is that architectural distortions are one of the most common findings in retrospective analyses of false-negative mammography and may represent the earliest manifestation of breast cancer [8]. The three types of lesions/abnormalities mentioned can be benign or malignant thus representing the absence or presence of breast cancer, respectively.

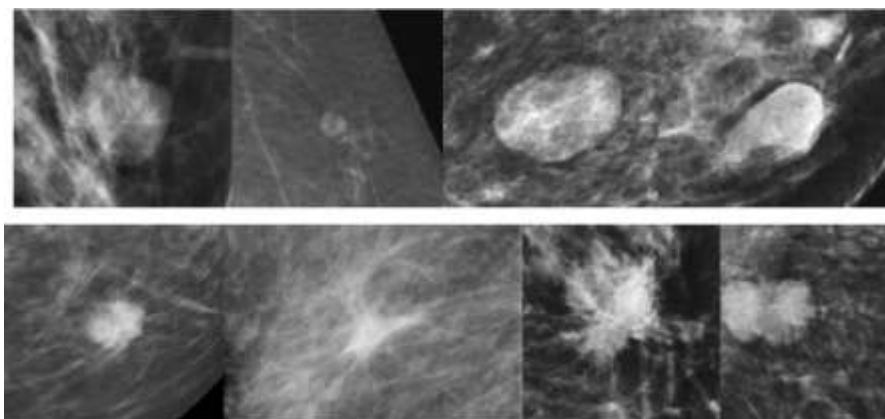


Figure 1.1. Examples of typically benign breast masses (up) and suspected malignant masses (down) [9, 10].

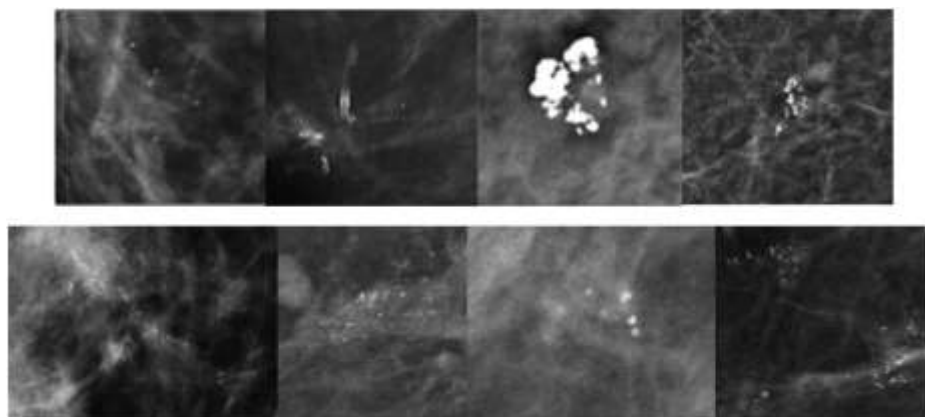


Figure 1.2. Examples of typically benign microcalcifications (up) and suspected malignant microcalcifications (down) (adapted from [11]).

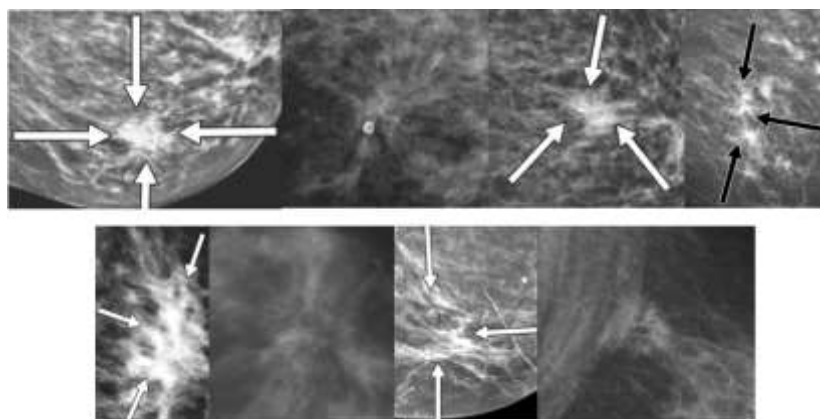


Figure 1.3. Examples of typically benign architectural distortions (up) and suspected malignant distortions (down) [8, 10, 12].

1.2. Digital Breast Tomosynthesis

Mammography is one of the most challenging technical areas of radiology as it requires high spatial resolution to capture small lesions such as microcalcifications and, simultaneously, high contrast to allow the detection of soft tissue lesions. In the 1970s, mammography was performed analogically, with the x-ray beams being captured on a radiographic film, which was then developed and reviewed by the physician. However, with the development of flat panel detectors, these films were replaced by converting information into a digital image displayed on high-resolution monitors [13]. This new approach, named full field Digital Mammography (DM), had its first equipment approved by the Food and Drug Administration (FDA) in 2000 [14] and since then it has been widely used in diagnosis of breast disease and as a screening modality.

Indeed, until recently, the gold standard for performing breast cancer screenings was DM. However, due to its two-dimensional (2D) acquisition nature that originates tissue overlap (Figure 1.4), it presents two recurrent limitations: (1) Lesions that are hidden by adjacent tissues, resulting in low sensitivity in dense breasts with pathology; and (2) normal regions that, when superimposed, resemble a lesion, resulting in low specificity. These two aspects raise several problems when using this technique. In many situations, when malignant lesions are not detected early enough as breast cancer, these lesions may spread, making the treatment much more complicated. In fact, missed cancers in DM are one of the most common reasons for malpractice lawsuits in radiology [15]. Conversely, the high rate of false positives causes a huge number of women to be recalled for complementary exams and/or biopsies that, in such cases, are completely unnecessary, causing anxiety for patients and additional financial costs [16].

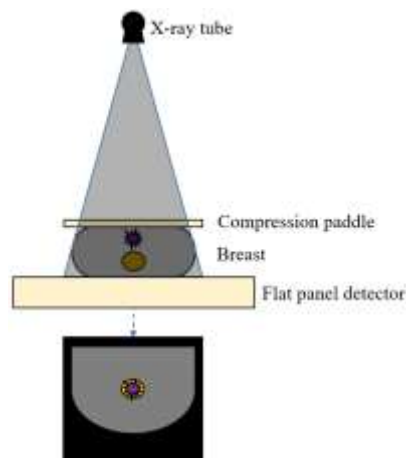


Figure 1.4. 2D DM Acquisition schematic.

In 2011, a technique that combines the concepts of DM and Computed Tomography (CT) was approved by the FDA [17]. This technique is Digital Breast Tomosynthesis (DBT) and it has an acquisition geometry similar to DM. With the breast compressed (such as in DM), the main difference is that in DBT the x-ray tube rotates around the breast and a set of low-dose 2D projections are acquired over a limited angular range allowing a “quasi” three-dimensional (3D) acquisition of the breast tissue (unlike CT, which acquires over 360°). The projection data are then reconstructed to produce several slice images of the breast tissue (Figure 1.5) and the depth information is added, addressing the limitations imposed by the 2D nature of DM. As a result, DBT improves the perception of the location and shape of lesions, without a significant increase of the radiation dose to the patient, when compared to 2D DM [18].

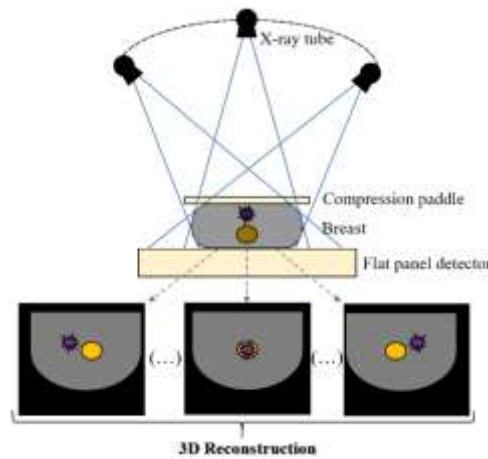


Figure 1.5. 3D DBT Acquisition schematic.

The number of acquired projections, as well as the angular range and the image reconstruction algorithm may vary by manufacturer, as shown in the Table 1.1.

Table 1.1. Some characteristics of FDA approved DBT systems [19].

	Hologic Selenia Dimensions	GE SenoClaire/ Sengraphe	Siemens MAMMOMAT Inspiration	Fujifilm ASPIRE
No. of projections	15	9	25	15
Angular range	$\pm 7.5^\circ$	$\pm 12.5^\circ$	$\pm 25^\circ$	$\pm 7.5^\circ / \pm 20^\circ$
X-ray tube motion	Continuous	Step and shoot	Continuous	Continuous
Reconstruction algorithm	FBP	ASiR	FBP	FBP

FBP: Filtered Back Projection / ASiR: Adaptive Statistical iterative Reconstruction

Initially, DBT was studied and approved in combination with DM, demonstrating an increase in breast cancer detection rates and a significant reduction in recall rates [20-25], particularly

with dense breasts [22]. Currently, by including synthetic mammography generated from DBT data, DBT alone is accepted as a stand-alone modality to replace DM [26-32]. In Figure 1.6 and Figure 1.7 are two examples of how DBT images provided a more accurate diagnosis.

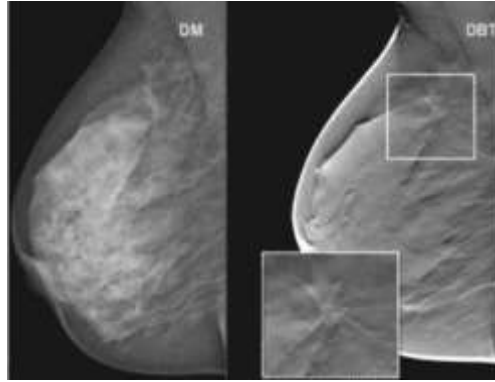


Figure 1.6. Example of a cancer detected by DBT alone and missed with DM [28].

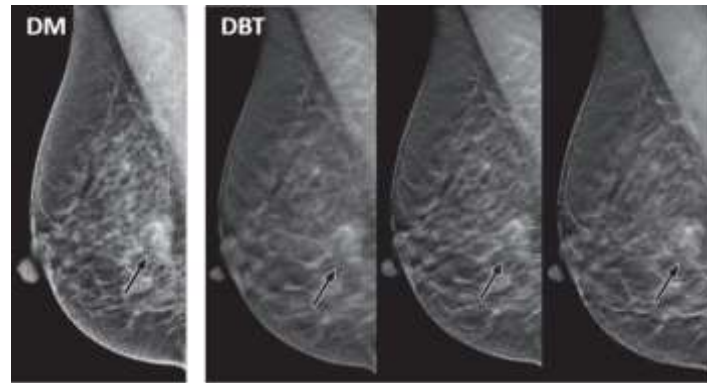


Figure 1.7. Example showing that structures at different levels in the breast can summate to create suspicious region on DM that may be identified as negative or superimposed tissue on DBT images. Left: DM with a suspicious area (arrow); right: three slices of DBT with no suspicious area [23].

1.3. Visualization of DBT

The 3D visualization of DBT is one of the most important and crucial aspects to correctly extract all the information provided by this technique. So far, DBT images are analyzed with a 2D slice-by-slice visualization or sequentially as a continuous cine loop [33]. This procedure leads to a subjective perception and a qualitative judgment since a 2D visualization of 3D data only allows an immediate inspection of anatomical structures in two dimensions of the plane, while the third dimension is mentally reconstructed looking at adjacent slices [34]. This can make the judgment of potential lesions, such as microcalcification clusters (MCs), difficult since MCs can be spread across several slices and may be difficult to interpret in a 2D visualization [35]. In addition, DBT presents, on average, sixty slices per exam (about thirty times more images than 2D DM), which results in time-consuming analysis both in screening and daily clinical use [36-38], directly influencing the number of interpreted examinations [38].

Currently, two other new approaches have emerged for the visualization of DBT data: synthetic mammography (SM) build upon DBT data [30, 39] and thicker slabs obtained by combining several slices [40, 41]. The existence of SM is very important because it allows a fair comparison with previous DM examinations (for retrospective analysis) and, in some situations, SM could benefit from the Computer Aided Detection/Diagnosis (CAD) systems developed for 2D application. However, SM still presents the same disadvantages as 2D DM (tissue overlapping) [42]. On the other hand, a system that uses artificial intelligence (AI) to reduce the number of slices to be analyzed and is based on thicker 6 mm slices (slabs) combined with synthesized 2D information was recently approved by the FDA [43, 44]. The construction of thicker slabs has demonstrated good results in terms of reducing time and false positives but originates a lower sensitivity, which can reach about 10% inferiority compared to the standard protocol that includes observation by planes [41].

A different type of visualization may play an important complementary role in breast cancer diagnosis [45]. Volume rendering (VR) is the process of creating realistic computer-generated scenes of volumetric data, yielding a true depth perception. VR can display data from any angle, resulting in an immediate global inspection and providing an understanding of the underlying data at once [34]. This type of supplementary visualization is already used in other tomographic medical imaging modalities such as CT, Magnetic Resonance Imaging, or Positron Emission Tomography [46-48]. There are some works mentioning a few aspects of 3D VR for DBT [49-51] and its importance to detect MCs [52]. As in CT, the integration of a 3D visualization in a 2D reading can be seen as a problem solver in some situations (for example, the need to better understand the shape of a MC) [53].

Ray casting is the classic algorithm used for VR, where rays are cast from the eye or other viewpoints and traverse a scene containing a volumetric data set. This process encompasses several parameters that define the appearance of the final rendered image. The ray function and the transfer functions are two of the most critical rendering parameters. In fact, the basic idea of VR can be divided into two interconnected phases that rely on these parameters: (1) a ray is projected through the data and the value of each voxel is determined according to a predefined ray function; and (2) this value is assigned to a specific color or opacity through the transfer functions.

Maximum Intensity Projection (MIP) and the compositing technique are the most commonly used ray functions for rendering medical data sets. These techniques evaluate the data along a line from the viewer's eye through the data set in a different way. With the MIP technique, each

voxel is evaluated and the maximum value is selected (Figure 1.8 (a)). As the displayed intensity will represent only the material with the highest intensity along the projected ray, a high intensity material, for example skin or calcifications may obscure information [54]. On the other hand, the compositing technique considers all the values of the data along that line, sums the contribution of each voxel and displays the resulting composite for each voxel (Figure 1.8 (b)). Two practical examples of Magnetic Resonance data rendered using each ray function are shown in Figure 1.9.

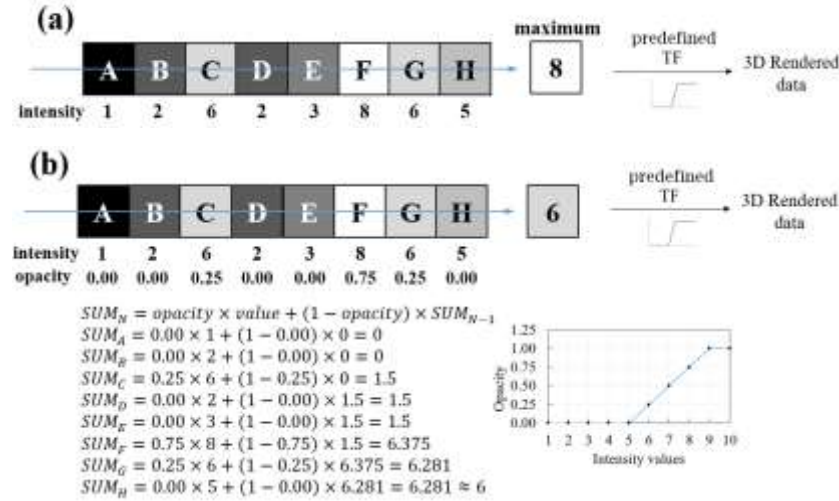


Figure 1.8. Illustrative scheme of MIP (a) and compositing (b) technique [54].

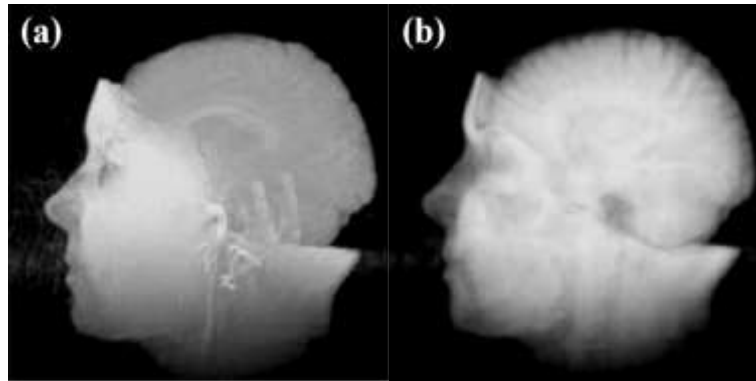


Figure 1.9. Examples of magnetic resonance rendered images obtained with (a) MIP and (b) the compositing technique [55].

Transfer functions are other fundamental parameter in VR and are particularly important to the quality of volume rendered images because they define how much and which data are visible by assigning optical properties like color and opacity to the voxel data [54]. These voxel data may refer to one (scalar intensity values) or two components (scalar intensity and gradient magnitude values) [56]. 1D transfer functions are based only on scalar intensity values and are the most commonly used form of transfer function. For medical image data, the 1D transfer function is often inadequate as tissues are represented by similar and overlapping intensities

[57]. In addition, DBT is very noisy, which may further reduce the ability of 1D transfer functions to correctly classify different tissue types [58]. If the transfer function takes into account more than one quantity on its domain, it is termed a 2D transfer function [58]. For medical imaging in general and for breast imaging in particular, a useful second dimension is the gradient magnitude because it measures how quickly values are changing in the image space. In this way, introducing it in the domain of a transfer function allows some discrimination between homogeneous and border regions [59]. Since tissues are relatively homogeneous, their gradient magnitudes are low and the peaks in the gradient histograms correspond to intensity changes, which means there are boundaries between two materials [60].

Due to the complexity of DBT data, the compositing technique becomes useful if used with the information extracted from the gradient. As it considers that all data contributes to the VR (not only the higher intensity, as in the case of MIP), it accumulates more information, which can only be distinguished with the gradient information.

Finding proper transfer functions is extremely difficult and is one of the major challenges in volume visualization. For example, in breast data, although a feature of interest may be easily identifiable in the spatial domain, the range of intensity values that characterize the feature are difficult to isolate in the transfer function domain. This results from the fact that other regions (not related to the feature of interest) contain the same range of intensities [60]. Even if the only variable which needs to be set is intensity (1D Transfer functions), it is time consuming and accomplished by a frustrating trial and error process [56, 59, 60]. Automatic and semi-automatic generation of transfer functions is the ultimate goal in many applications since they enable a more widespread use of VR [56, 61, 62].

1.4. Automatic classification and detection of lesions

In a screening environment, only a small part of the exams are representative of cancer, meaning that there are many more normal screens than those with pathology. For example, in Europe, in 1000 screens, only about 6.4 and 8.8 cancer cases are detected with DM and DBT, respectively [63]. In other words, among so many "normal" slices that the radiologist analyzes, there will be only a few that represent breast cancer lesions (sometimes with a very subtle presentation) that will be flagged. For this reason, automatic or semi-automatic computer systems that contribute to the detection and/or classification of potential lesions are extremely important, especially in a screening environment.

1.4.1. CAD systems

CAD systems are intended to assist radiologists in this arduous task. The CAD software for screening mammography was first approved by the FDA in 1998 [64] and has been widely adopted in clinical practice [65]. In general, these systems place markers in suspicious areas that are later evaluated by the radiologists, thus alleviating the task of detecting/finding tiny spots in the middle of many exams. If they are responsible for an initial detection and localization of suspicious lesions (benign or malignant), they are called Computer Aided Detection (CADe). If, on the other hand, their function is to interpret suspicious regions, assigning for example a score to the level of malignancy, they are called Computer Aided Diagnosis (CADx).

DBT CAD systems were implemented and evaluated in an attempt to shorten the reading time while maintaining the radiologist performance. In fact, some results were very encouraging with reading time reductions between 14% and 29.2% without loss of diagnostic performance [66-68]. However, despite the efforts and improvements already achieved (such as decreasing the false negative rate), these CAD systems have not reached a level of performance that can be translated into a true improvement in the real screening of breast cancer due to the high false positive rates and low specificity [69-72].

1.4.2. Deep Convolutional Neural Networks (CNNs)

The human brain is considered the best organized system for efficient processing of information obtained from different senses. One of the mechanisms used in this processing is the treatment of high-level information through the collaboration/connection of a large number of simple structural elements (neurons). In machine learning, artificial neural networks are complex hierarchical models that intended to mimic the thought processes in the human brain and learn patterns present in observations [73]. Since 2012, with the success of deep convolutional neural networks (CNNs) in the ImageNet Large Scale Visual Recognition Challenge [74], the advances in computational power (with the wide availability of graphics processing units (GPU)) and the introduction of large scale datasets, there has been a new interest in developing better automatic methods of image analysis in general and medical imaging in particular.

CNNs are used to analyze images because the structural information between neighboring pixels or voxels is a crucial source of information in those cases. CNNs include a set of filters (convolutional kernels) that are convolved with a given input to generate an output feature map. These complex features are used to identify essential information that is needed to generate the

final output. When multiple layers are used, we refer to deep CNNs. In addition to convolutional layers, typical CNNs consist of many other layers: input layers, Rectified Linear Unit (ReLU) layers (also referred as activation function, which performs a threshold operation in each element of the input where any value less than zero is set to zero), normalization layers, dropout layers (randomly set input elements to zero with a given probability), max/average pooling layers (perform downsampling by dividing the input into rectangular pooling regions and then computing the maximum/average of each region), fully connected layers (that generally comprise the last few layers of the CNN and that connects all neurons of the final max-pooling or convolution layer to each of the output neurons), output layers, etc.

A CNN typically follows sequential stages of convolutional layers combined with pooling layers which are then followed by fully connected layers (Figure 1.10). These models have the ability to decompose the image information into different representations from low to high level features. The layers near the input layer are more general and the layers next to the output layer are more specific regarding the source image [75].

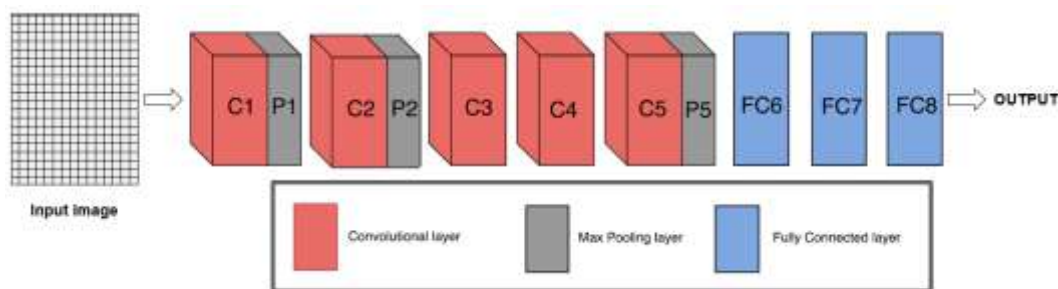


Figure 1.10. Illustration of the sequential arrangement of layers in a very simple CNN (adapted from [73]).

The models can be divided according to their main task: Classification models whose main function is to identify the probability of an image or part of the image belonging to a predicted class; Object detection models whose aim is to locate specific objects in an image; and segmentation models that label pixels or voxels individually as belonging or not to a certain class [76].

1.4.3. Learning

The CNN stores information in the form of weights and bias. A weight is assigned to each connection between two elements of two consecutive layers, which means that the weight assigned to each feature will decide its influence on the output prediction. The bias are constants added to the product of inputs and weights, used to give more flexibility to the models, shifting the activation function towards the positive or negative side. It is important to have a powerful model but also an algorithm to learn the model's parameters (weights and bias) efficiently.

The learning process is performed through optimization algorithms to minimize the loss function that represents the error of the model. Gradient descent requires calculating the gradient of the loss function for each variable in the model and is probably the most used minimization algorithm in deep learning [77]. The calculation of this gradient is efficiently performed through an algorithm called back-propagation that uses the chain rule to speed up the computation [73, 77]. Together, the gradient descent and the back-propagation algorithms are usually used to train a neural network, i.e., to adjust the weights to reduce the error.

This parameter update can be done based on the gradients evaluated over all the training samples (batch gradient descent) or sequentially by calculating the gradient based on one sample at a time (stochastic gradient descent). Typically, to have a trade-off between these two approaches, the computation and update of the parameters are performed by stochastic gradient descent with a mini-batch that contains a small set of samples, rather than a single sample [73]. In this case, each parameter update is called an iteration and a full pass through the entire data set is called an epoch.

1.4.4. CNNs in breast cancer detection

Recently, there have been several large-scale published studies whose aim was to analyze AI systems based on CNNs as a potential solution in the field of automated breast cancer in DM and DBT [72]. In some cases, the performance of the AI system has been compared to that of the breast radiologist alone, achieving a comparable or even improved cancer detection accuracy [78-80]. In other studies, where the performance of the radiologist with and without AI was evaluated, physicians improved their cancer detection when using an AI system for support [68, 80-83]. With these promising results and the need for an automatic detection system for lesions in DBT and in screening, much research is being carried out in this regard.

1.5. Thesis outline

The main objective of this thesis is to answer to some daily clinical challenges related to the visualization of DBT data, as well as lesion detection. For that, new computational imaging methods are developed to increase the diagnostic value of DBT by helping the radiologists in their task of reading DBT volumes and diagnose breast cancer. Different methods involving visualization and automatic lesion detection are studied. These methods can potentially aid in the early detection of breast cancer and speed up the process of assessing DBT volumes, allowing an inspection of the entire dataset at once.

The thesis is composed of eight chapters. In addition to the introduction (chapter 1) and the final considerations (chapter 8), this thesis consists of six main chapters that describe the developed work. Chapters 2, 3, 4 and 5 include methodology related to the VR visualization of DBT data and improvement of rendered images, and chapters 6 and 7 are related to the automatic detection of lesions in DBT.

Chapter 1 consists of a general introduction, where the concepts of the work are described.

Chapter 2 describes a methodology to automatically generate transfer functions that result in appropriate and useful VR visualizations of DBT data. This function is probably the factor that most influences the rendered images in the sense that it defines what is made opaque and what is transparent. It is usually through a trial-and-error exploration of the transfer functions that the desired renderings attempt to be achieved. However, this is a time-consuming process and often does not directly result in the visualization of what is desired. In this way, the generation of transfer functions in an automatic or semi-automatic way is extremely important not only to somehow standardize these types of visualizations, but also to make this process more accessible to the radiologists, so that they can see in VR a complementary way of analyzing DBT data. This chapter was published as a conference paper in the renowned conference on breast imaging – the *15th International Workshop on Breast Imaging (IWBI2020)*.

Chapter 3 presents the study of the influence of two rendering parameters (sampling distance and voxel dimension in z direction) on DBT rendered data. Both the parameters have a significant impact on the quality of the final rendered volume. If the sampling distance is too large, our sampling might miss important features in the data and generate major aliasing artifacts. Alternatively, if a very small distance is considered, the number of samples collected along the ray is increased and the amount of time required to render the image will significantly increase. On the other hand, reconstructed voxels in DBT are generally anisotropic, with the size in z direction greater than the size in x and y directions. With the visualization slice-by-slice, this does not have much influence on the quality since the data is observed in xy plane. However, with rendering, the volume of DBT data can be analyzed from any angle, making the z -voxel dimension important as well. In this way, several interpolation methods and their properties to make the voxel isotropic are studied. Taking into consideration these two parameters, a balance between quality and time is achieved by analyzing several figures of merit. This chapter was published as a journal paper in the *Journal of Imaging*. This journal is Q2 in “Computer Graphics and Computer-Aided Design”, “Computer Vision and Pattern

Recognition” and “Electrical and Electronic Engineering” domains and Q3 in “Radiology, Nuclear Medicine and Imaging” field.

Chapter 4 addresses the implementation and evaluation of two total variation minimization methods in order to decrease noise while preserving edges. In DBT, both the low dose per projection and the angular limit raise some complications, such as the high noise level in the reconstructed images. The first algorithm described is a 2D total variation minimization filter sequentially applied to all slices, one slice at the time, while the second is a 3D TV minimization filter applied to the entire volume. Their performance is analyzed through 3D visualization of DBT data with VR. This chapter was published as a journal paper in the *Computer Methods and Programs in Biomedicine*. This journal has an impact factor of 5.428 and is ranked in the Q1 in the domains: “Computer science – interdisciplinary applications”, “Computer science – theory & methods”, “Biomedical Engineering” and “Medical Informatics”.

Chapter 5 presents a novel approach in DBT through the study of out-of-plane artifacts using blind deconvolution and noise regularization based on total variation minimization. DBT presents out-of-plane artifacts (blur in the z-direction) caused by features of high intensity (calcifications, biopsy needles and localization wires). The production of these artifacts could potentially obscure breast lesions and would limit the usability of DBT in interventional procedures. In this way, four different approaches to blind deconvolution are presented and the results are analyzed using real phantom data under conventional slice-by-slice visualization and 3D VR with a compositing technique. This chapter was published as a journal paper in the *IEEE Transactions on Medical Imaging*. This journal have an impact factor of 10.048 and is ranked in the Q1 in the domains: “Computer science – interdisciplinary applications”, “Biomedical Engineering”, “Electrical & Electronic Engineering”, “Imaging science & photographic technology” and “Radiology, Nuclear medicine & Medical imaging”.

Chapter 6 introduces fully automatic methods based on deep CNNs in order to study the classification of DBT images about the presence or absence of MCs. One of the great difficulties in training AI algorithms with DBT data is the lack of labelled DBT databases. Furthermore, all published studies refer to private databases. Using a public database with synthetic DBT data of virtual patients, the aim of this part of the work is to input a whole DBT image and have a direct answer about the absence or presence of MCs, without the need for prior identification of lesions in specific regions. Four popular deep CNNs are trained and compared with one new architecture proposed in this thesis. Besides the original data, six different pre-processing meth-

odologies, whose main purpose was to highlight microcalcifications, are implemented to generate different input datasets. In this chapter, an extensive review of the literature on classification/detection of lesions using CNNs in DBT is described. This chapter was accepted under revision for publication as a journal paper in the *Medical Physics*, which have an impact factor of 4.071 and is ranked in the Q1 in the domain: “Radiology, Nuclear medicine & Medical imaging”.

Chapter 7 is the link between the automatic detection of MCs in DBT and the visualization of these small lesions using the techniques developed for VR. Through the training of a faster region-based CNN (faster R-CNN) and, with the knowledge obtained from the work in chapter 6, the MCs are located through predicted bounding boxes and their respective scores. This information is then introduced into the visualization software and, instead of the entire breast being rendered, it is possible to obtain a 3D region of interest with magnification of these flagged MCs. This 3D zoom can help radiologists in the interpretation of these lesions since (1) it indicates specific regions of the breast with a possible lesion that deserve additional attention and (2) as the rendering of the MCs is almost like a segmentation of them, a detailed complementary analysis of their morphology is possible. This chapter was accepted for oral presentation and publication as an extended conference paper in the *BioImaging2022* conference. BioImaging is part of BIOSTEC, the International Joint Conference on Biomedical Engineering Systems and Technologies.

Chapter 8 consists of the final considerations and includes a summary and general discussion of each development chapter, the main limitations of this thesis and future work.

References

1. Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries*. CA Cancer J Clin, 2021. **71**(3): p. 209-249.
2. Siegel, R.L., et al., *Cancer Statistics, 2021*. CA Cancer J Clin, 2021. **71**(1): p. 7-33.
3. Tabár, L., et al., *The incidence of fatal breast cancer measures the increased effectiveness of therapy in women participating in mammography screening*. Cancer, 2019. **125**(4): p. 515-523.
4. Peintinger, F., *National Breast Screening Programs across Europe*. Breast Care, 2019. **14**(6): p. 354-358.
5. D'Orsi, C.J., *2013 ACR BI-RADS Atlas: Breast Imaging Reporting and Data System*2014: American College of Radiology.
6. Azam, S., et al., *Mammographic microcalcifications and risk of breast cancer*. British journal of cancer, 2021. **125**(5): p. 759-765.
7. Cox, R.F., et al., *Microcalcifications in breast cancer: novel insights into the molecular mechanism and functional consequence of mammary mineralisation*. British journal of cancer, 2012. **106**(3): p. 525-537.

8. Gaur, S., et al., *Architectural Distortion of the Breast*. American Journal of Roentgenology, 2013. **201**(5): p. W662-W670.
9. Berment, H., et al., *Masses in mammography: What are the underlying anatomopathological lesions?* Diagnostic and Interventional Imaging, 2014. **95**(2): p. 124-133.
10. Selvi, R., *Breast diseases: imaging and clinical management* 2014: Springer.
11. Wilkinson, L., V. Thomas, and N. Sharma, *Microcalcification on mammography: approaches to interpretation and biopsy*. The British Journal of Radiology, 2017. **90**(1069): p. 20160594.
12. Barter, S., *Carcinoma, Breast, Imaging Mammography, Primary Signs*, in *Encyclopedia of Diagnostic Imaging*, A.L. Baert, Editor 2008, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 220-224.
13. Zackrisson, S. and N. Houssami, *Chapter 13 - Evolution of Mammography Screening: From Film Screen to Digital Breast Tomosynthesis*, in *Breast Cancer Screening*, N. Houssami and D. Miglioretti, Editors. 2016, Academic Press: Boston. p. 323-346.
14. Food and Drug Administration (FDA) U.S. . *Premarket Approval GE Healthcare Senographe Full Field Digital Mammography system*. 2000 [cited 2021 October]; Available from: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P990066>.
15. Whang, J.S., et al., *The Causes of Medical Malpractice Suits against Radiologists in the United States*. Radiology, 2013. **266**(2): p. 548-554.
16. Tosteson, A.N.A., et al., *Consequences of False-Positive Screening Mammograms*. JAMA Internal Medicine, 2014. **174**(6): p. 954-961.
17. Food and Drug Administration (FDA) U.S. . *Accreditation/Certification for Facilities Utilizing a DBT System with FFDM Images*. 2011 [cited 2020 February]; Available from: <http://www.fda.gov/Radiation-EmittingProducts/MammographyQualityStandardsActandProgram/FacilityCertificationandInspection/ucm114148.htm>.
18. Svahn, T.M., et al., *Review of radiation dose estimates in digital breast tomosynthesis relative to those in two-view full-field digital mammography*. The Breast, 2015. **24**(2): p. 93-99.
19. Tirada, N., et al., *Digital Breast Tomosynthesis: Physics, Artifacts, and Quality Control Considerations*. Radiographics, 2019. **39**(2): p. 413-426.
20. Ciatto, S., et al., *Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study*. The Lancet Oncology, 2013. **14**(7): p. 583-589.
21. Skaane, P., et al., *Performance of breast cancer screening using digital breast tomosynthesis: results from the prospective population-based Oslo Tomosynthesis Screening Trial*. Breast Cancer Research and Treatment, 2018. **169**(3): p. 489-496.
22. Haas, B.M., et al., *Comparison of Tomosynthesis Plus Digital Mammography and Digital Mammography Alone for Breast Cancer Screening*. Radiology, 2013. **269**(3): p. 694-700.
23. Rose, S.L., et al., *Implementation of Breast Tomosynthesis in a Routine Screening Practice: An Observational Study*. American Journal of Roentgenology, 2013. **200**(6): p. 1401-1408.
24. Greenberg, J.S., et al., *Clinical Performance Metrics of 3D Digital Breast Tomosynthesis Compared With 2D Digital Mammography for Breast Cancer Screening in Community Practice*. American Journal of Roentgenology, 2014. **203**(3): p. 687-693.
25. McDonald, E.S., et al., *Effectiveness of Digital Breast Tomosynthesis Compared With Digital Mammography: Outcomes Analysis From 3 Years of Breast Cancer Screening*. JAMA Oncology, 2016. **2**(6): p. 737-743.
26. Zackrisson, S., et al., *One-view breast tomosynthesis versus two-view mammography in the Malmö Breast Tomosynthesis Screening Trial (MBTST): a prospective, population-based, diagnostic accuracy study*. The Lancet Oncology, 2018. **19**(11): p. 1493-1503.
27. Bernardi, D., et al., *Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study*. The Lancet Oncology, 2016. **17**(8): p. 1105-1113.
28. Lång, K., et al., *Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmö Breast Tomosynthesis Screening Trial, a population-based study*. Eur Radiol, 2016. **26**(1): p. 184-190.

29. Gilbert, F.J., et al., *Accuracy of Digital Breast Tomosynthesis for Depicting Breast Cancer Subgroups in a UK Retrospective Reading Study (TOMMY Trial)*. Radiology, 2015. **277**(3): p. 697-706.
30. Hofvind, S., et al., *Digital Breast Tomosynthesis and Synthetic 2D Mammography versus Digital Mammography: Evaluation in a Population-based Screening Program*. Radiology, 2018. **287**(3): p. 787-794.
31. Freer, P.E., et al., *Clinical implementation of synthesized mammography with digital breast tomosynthesis in a routine clinical practice*. Breast Cancer Research and Treatment, 2017. **166**(2): p. 501-509.
32. Food and Drug Administration (FDA) U.S. . *Premarket Approval application supplement for the Selenia Dimensions 3D System*. 2013 [cited 2021 May].
33. Sechopoulos, I., *A review of breast tomosynthesis. Part I. The image acquisition process*. Med Phys, 2013. **40**(1): p. 014301.
34. Suetens, P., *Medical image analysis*, in *Fundamentals of Medical Imaging* 2009, Cambridge University Press: New York. p. 159-189.
35. Samala, R.K., et al., *Digital breast tomosynthesis: Computer-aided detection of clustered microcalcifications on planar projection images*. Phys Med Biol, 2014. **59**(23): p. 7457-7477.
36. Good, W.F., et al., *Digital breast tomosynthesis: a pilot observer study*. AJR Am J Roentgenol, 2008. **190**(4): p. 865-9.
37. Gur, D., et al., *Digital breast tomosynthesis: observer performance study*. AJR Am J Roentgenol, 2009. **193**(2): p. 586-91.
38. Caumo, F., et al., *Digital Breast Tomosynthesis with Synthesized Two-Dimensional Images versus Full-Field Digital Mammography for Population Screening: Outcomes from the Verona Screening Program*. Radiology, 2018. **287**(1): p. 37-46.
39. Simon, K., et al., *Accuracy of Synthetic 2D Mammography Compared With Conventional 2D Digital Mammography Obtained With 3D Tomosynthesis*. American Journal of Roentgenology, 2019. **212**(6): p. 1406-1411.
40. van Schie, G., et al., *Mass detection in reconstructed digital breast tomosynthesis volumes with a computer-aided detection system trained on 2D mammograms*. Med Phys, 2013. **40**(4): p. 041902.
41. Iotti, V., et al., *Comparing two visualization protocols for tomosynthesis in screening: specificity and sensitivity of slabs versus planes plus slabs*. Eur Radiol, 2019. **29**(7): p. 3802-3811.
42. Petropoulos, A.E., et al., *Quantitative assessment of microcalcification cluster image quality in digital breast tomosynthesis, 2-dimensional and synthetic mammography*. Medical & Biological Engineering & Computing, 2020. **58**(1): p. 187-209.
43. Food and Drug Administration (FDA) U.S. . *Approval for software option 3DQuorum™ technology - Premarket Approval*. 2019 [cited 2020 25 June]; Available from: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P080003S008>.
44. *3DQuorum™ Imaging Technology - Improving radiologist performance through Artificial Intelligence and SmartSlices (White Paper)*. 2020; Available from: [https://www.hologic.com/sites/default/files/downloads/WP-00152_Rev001_3DQuorum_Imaging_Technology_Whitepaper%20%20\(1\).pdf](https://www.hologic.com/sites/default/files/downloads/WP-00152_Rev001_3DQuorum_Imaging_Technology_Whitepaper%20%20(1).pdf).
45. Venson, J.E., et al., *A Case-Based Study with Radiologists Performing Diagnosis Tasks in Virtual Reality*. Stud Health Technol Inform., 2017. **245**: p. 244-248.
46. O'Connell, A., et al., *Cone-Beam CT for Breast Imaging: Radiation Dose, Breast Coverage, and Image Quality*. American Journal of Roentgenology, 2010. **195**(2): p. 496-509.
47. Song, H., X. Cui, and F. Sun, *Breast Tissue 3D Segmentation and Visualization on MRI*. International Journal of Biomedical Imaging, 2013. **2013**: p. 8.
48. Jung, Y., et al., *Occlusion and Slice-Based Volume Rendering Augmentation for PET-CT*. IEEE Journal of Biomedical and Health Informatics, 2017. **21**(4): p. 1005-1014.
49. Alyassin, A.M. *Automatic transfer function generation for volume rendering of high-resolution x-ray 3D digital mammography images*. in *Medical Imaging 2002*. 2002. SPIE.
50. Alyassin, A.M., et al., *3D Visualization of X-ray Tomosynthesis Digital Mammography Data: Preference Study*, in *Digital Mammography: IWDM 2002 — 6th International Workshop on Digital Mammography*, H.-O. Peitgen, Editor 2003, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 507-509.

51. Dharanija, R. and T. Rajalakshmi. *A Conjunct Analysis for Breast Cancer Detection by Volume Rendering of Low Dosage Three Dimensional Mammogram*. in *Progress In Electromagnetics Research Symposium Proceedings*. 2011. China.
52. Jerebko, A., et al. *3D rendering methods for visualization of clusters of calcifications in digital breast tomosynthesis: a feasibility study*. in *ECR 2011*. 2011. Vienna, Austria.
53. Neri, E., et al., *The second ESGAR consensus statement on CT colonography*. *Eur Radiol*, 2013. **23**(3): p. 720-729.
54. Calhoun, P.S., et al., *Three-dimensional Volume Rendering of Spiral CT Data: Theory and Method*. *Radiographics*, 1999. **19**(3): p. 745-764.
55. Schroeder, W., K. Martin, and B. Lorensen, *The Visualization Toolkit: An Object-oriented Approach to 3D Graphics*. 4rd ed2006, USA: Kitware.
56. Kindlmann, G. and J.W. Durkin. *Semi-automatic generation of transfer functions for direct volume rendering*. in *IEEE Symposium on Volume Visualization (Cat. No.989EX300)*. 1998.
57. Lundstrom, C., P. Ljung, and A. Ynnerman, *Local Histograms for Design of Transfer Functions in Direct Volume Rendering*. *IEEE Transactions on Visualization and Computer Graphics*, 2006. **12**(6): p. 1570-1579.
58. Ljung, P., et al., *State of the Art in Transfer Functions for Direct Volume Rendering*. *Computer Graphics Forum*, 2016. **35**(3): p. 669-691.
59. Kindlmann, G. *Transfer functions in direct volume rendering: Design, Interface, Interaction*. September 2017 November 2019]; Available from: <http://www.cs.utah.edu/~gk/papers/sig02-TF-notes.pdf>.
60. Kniss, J.K., G.; Hansen, C., *Visualization Handbook*, 2004, Academic Press. p. 181-202.
61. Correa, C.D. and K.L. Ma, *Visibility Histograms and Visibility-Driven Transfer Functions*. *IEEE Transactions on Visualization and Computer Graphics*, 2011. **17**(2): p. 192-204.
62. Ma, B. and A. Entezari, *Volumetric Feature-Based Classification and Visibility Analysis for Transfer Function Design*. *IEEE Transactions on Visualization and Computer Graphics*, 2017. **PP**(99): p. 1-1.
63. Marinovich, M.L., et al., *Breast Cancer Screening Using Tomosynthesis or Mammography: A Meta-analysis of Cancer Detection and Recall*. *JNCI: Journal of the National Cancer Institute*, 2018. **110**(9): p. 942-949.
64. Food and Drug Administration (FDA) U.S. . *M1000 ImageChecker*. 1998 [cited 2021 May]; Available from: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?ID=319829>.
65. Keen, J.D., J.M. Keen, and J.E. Keen, *Utilization of Computer-Aided Detection for Digital Screening Mammography in the United States, 2008 to 2016*. *Journal of the American College of Radiology*, 2018. **15**(1, Part A): p. 44-48.
66. Balleyguier, C., et al., *Improving digital breast tomosynthesis reading time: A pilot multi-reader, multi-case study using concurrent Computer-Aided Detection (CAD)*. *Eur J Radiol*, 2017. **97**: p. 83-89.
67. Benedikt, R.A., et al., *Concurrent Computer-Aided Detection Improves Reading Time of Digital Breast Tomosynthesis and Maintains Interpretation Performance in a Multireader Multicase Study*. *American Journal of Roentgenology*, 2017. **210**(3): p. 685-694.
68. Chae, E.Y., et al., *Decrease in interpretation time for both novice and experienced readers using a concurrent computer-aided detection system for digital breast tomosynthesis*. *Eur Radiol*, 2019. **29**(5): p. 2518-2525.
69. Fenton, J.J., et al., *Influence of Computer-Aided Detection on Performance of Screening Mammography*. *New England Journal of Medicine*, 2007. **356**(14): p. 1399-1409.
70. Lehman, C.D., et al., *Diagnostic Accuracy of Digital Screening Mammography With and Without Computer-Aided Detection*. *JAMA Internal Medicine*, 2015. **175**(11): p. 1828-1837.
71. Katzen, J. and K. Dodelzon, *A review of computer aided detection in mammography*. *Clinical Imaging*, 2018. **52**: p. 305-309.
72. Sechopoulos, I., J. Teuwen, and R. Mann, *Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art*. *Seminars in Cancer Biology*, 2020.

73. Zhou, K., H. Greenspan, and D. Shen, *Deep Learning for Medical Image Analysis* 2017: Elsevier Science.
74. Krizhevsky, A., I. Sutskever, and G.E. Hinton, *Imagenet classification with deep convolutional neural networks*. Advances in neural information processing systems, 2012. **25**: p. 1097-1105.
75. Yosinski, J., et al., *How transferable are features in deep neural networks?*, in *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2* 2014, MIT Press: Montreal, Canada. p. 3320–3328.
76. Bai, J., et al., *Applying deep learning in digital breast tomosynthesis for automatic breast cancer detection: A review*. Med Image Anal, 2021. **71**: p. 102049.
77. Goodfellow, I., Y. Bengio, and A. Courville, *Deep Learning* 2016: MIT Press.
78. Rodríguez-Ruiz, A., et al., *Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists*. Journal of the National Cancer Institute, 2019. **111**(9): p. 916-922.
79. Kim, H.-E., et al., *Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study*. The Lancet Digital Health, 2020. **2**(3): p. e138-e148.
80. van Winkel, S.L., et al., *Impact of artificial intelligence support on accuracy and reading time in breast tomosynthesis image interpretation: a multi-reader multi-case study*. Eur Radiol, 2021.
81. Schaffter, T., et al., *Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms*. JAMA Network Open, 2020. **3**(3): p. e200265-e200265.
82. Rodríguez-Ruiz, A., et al., *Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System*. Radiology, 2019. **290**(2): p. 305-314.
83. Conant, E.F., et al., *Improving Accuracy and Efficiency with Concurrent Use of Artificial Intelligence for Digital Breast Tomosynthesis*. Radiology: Artificial Intelligence, 2019. **1**(4): p. e180096.

Calculation of transfer functions for volume rendering of breast tomosynthesis imaging

2

Ana M. Mota^{*a}, Matthew J. Clarkson^b, Lurdes Orvalho^c, Pedro Almeida^a and Nuno Matela^a,
"Calculation of transfer functions for volume rendering of breast tomosynthesis imaging," *Proc.
SPIE 11513, 15th International Workshop on Breast Imaging (IWBI2020)*, 1151327 (22 May
2020); doi: 10.1117/12.2559932

IWBI is a peer-reviewed conference and is one of the most prestigious international conferences on breast imaging.

^aInstituto de Biofísica e Engenharia Biomédica, Faculdade de Ciências, Universidade de Lisboa, Portugal (*ammota@fc.ul.pt);

^bDepartment of Medical Physics and Biomedical Engineering and the Centre for Medical Image Computing, University College London, London, UK;

^cHospital da Luz, Lisboa, Portugal

Abstract

Slice-by-slice visualization of Digital Breast Tomosynthesis (DBT) data is time consuming and can hamper the interpretation of lesions such as clusters of microcalcifications. With a visualization of the object through multiple angles, 3D volume rendering (VR) provides an intuitive understanding of the underlying data at once. 3D VR may play an important complementary role in breast cancer diagnosis. Transfer functions (TFs) are a critical parameter in VR and finding good TFs is a major challenge. The purpose of this work is to study a methodology to automatically generate TFs that result in appropriate and useful VR visualizations of DBT data.

For intensity-based TFs, intensity histograms were used to study possible relationships between statistics and critical intensity values in DBT data. The mean of each histogram has proved to be a valid option to automatically calculate those critical values that define these functions. At this stage, eight visualizations were obtained by combining several opacity/color intensity-based functions. Considering the gradient, ten visualizations were obtained. Nine of the ten TFs were constructed considering the peaks of gradient magnitude histograms. The tenth function was a simple linear ramp. Finally, three intensity-based and three gradient-based functions were selected and simultaneously used. This resulted in nine final VR visualizations taking both information into account.

The studied approach allowed an automatic generation of opacity/color TFs based on scalar intensity and gradient magnitude histograms. In this way, the preliminary results obtained with this methodology are very encouraging about creating an adequate visualization of DBT data by VR.

Keywords: Digital breast tomosynthesis, volume rendering, 3D visualization, transfer function

2.1. Introduction

Digital Breast Tomosynthesis (DBT) has been consolidating its position as a technique to replace 2D Digital Mammography (DM) in both screening and clinical environment [1-4]. So far, DBT data is displayed one slice at a time or sequentially as a continuous cine loop. This procedure can lead to difficult interpretation of microcalcification clusters, which can be spread across several slices [5]. In addition, the mean reading time doubles for DBT compared with DM examination [6-8]. The time required for each DBT data set evaluation is crucial both in clinical and screening environments, directly influencing the number of examinations interpreted [8].

A different type of visualization, such as 3D volume rendering (VR), may play an important complementary role in breast cancer diagnosis [9]. 3D VR is the process of creating realistic computer-generated images of a 3D scene, yielding a true depth perception [10]. With a visualization of the object through multiple angles, one of the advantages of 3D VR is to provide an intuitive understanding of the underlying data at once. Transfer functions (TFs) are a critical parameter in VR. They define how much and which data are visible by assigning opacity and color to the intensity and/or gradient magnitude values. To provide useful information about the volume data, TFs must ensure a balance between what is made transparent and what is considered to be of interest and must be opaque. Finding good TFs is one of the major challenges in volume visualization. It is time consuming and accomplished by a frustrating trial and error process [11-13]. Automatic and semi-automatic generation of TFs is the ultimate goal in many applications since it enables a more widespread use of VR [11, 14, 15].

There are some works that mention a few aspects of 3D VR for DBT [16-18] and its importance to detect clusters of microcalcifications [19]. However, there are currently no established methods or conclusions on acceptable TFs to provide adequate 3D VR visualization of DBT data. For this reason, opacity and color values were assigned to the intensity and gradient magnitude values of the data [20] and an innovative methodology to automatically generate TFs which allow an adequate 3D visualization by VR of DBT data is presented.

2.2. Methods

2.2.1. Data visualization

Twenty DBT data sets of anonymous patients from a clinical facility database were selected (Hospital da Luz S.A., Lisbon, Portugal). The heterogeneity of the sample was taken into account: there are ten cases acquired with mediolateral oblique view and ten with craniocaudal view, 50% are from right breasts and 50% from left breasts. All breast density levels, as well as all diagnostic classifications (based on the criteria of the American College of Radiology's Breast Imaging Reporting and Data System – BI-RADS) are included in this study.

The visualization software was developed in C++ using the Visualization Toolkit library (VTK) version 7.1.0. [21]. For a better visualization quality, voxels were made isotropic (with dimensions $0.085 \times 0.085 \times 0.085 \text{ mm}^3$) using the Lanczos function available in VTK (used by default) [22]. 3D VR was obtained with composite technique.

2.2.2. Generation of scalar intensity-based TFs

For each of the twenty cases, two regions were selected and the respective intensity values recorded. Region 1 is related to adipose tissue (similar to background) and region 2 refers to high intensity materials, such as microcalcifications. As the twenty cases comprehend all different levels of BI-RADS diagnosis, in cases with calcifications, region 2 includes data both on benign and malignant calcifications.

$a1$ and $b1$ stand for the intensity value of region 1 and 2, respectively. Voxels with scalar intensity values below $a1$ are considered totally transparent, above $b1$ are totally opaque and with intermediate values correspond to a linear relationship between intensity and opacity/color (Figure 2.1).

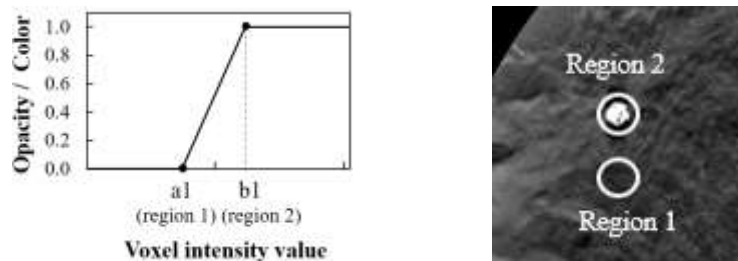


Figure 2.1. Left: 1D linear ramp transfer function used to map voxel intensity values to opacity and color. Right: Zoom in on the region 1 and region 2 whose intensities correspond to $a1$ and $b1$, respectively.

Three intensity histogram statistics – mean, maximum and skewness – were calculated for each case. The relationships between $a1$ and $b1$ and the intensity histogram statistics were studied to obtain possible expressions to estimate $a1$ and $b1$, having knowledge of the statistics. Additionally, two other intensity values – $a2$ and $b2$ – were calculated in order to achieve more distinct visualizations (Eq. 2.1 and Eq. 2.2). Taking into account different combinations of this critical values, eight visualizations were obtained.

$$a2 = a1 + \frac{b1 - a1}{2} \quad \text{Eq. 2.1}$$

$$b2 = b1 + \frac{b1 - a1}{2} \quad \text{Eq. 2.2}$$

2.2.3. Generation of gradient magnitude-based TFs

Gradient magnitude is a useful second dimension because it measures how quickly values are changing in the image space. Breast tissues are relatively homogeneous, so their gradient magnitudes are low. The peaks in the gradient histograms correspond to sudden intensity changes, which means, boundaries between two materials [13].

Here, to only have one degree of freedom (gradient), the influence of intensity-based TFs was kept neutral (opacity and color values were kept at one for all scalar intensity values). For each DBT case, three peaks in the gradient magnitude histogram were found. Based on these peaks, nine opacity TFs were analyzed. In addition, one function independent from the peaks, with a shape of a linear ramp crossing all magnitude gradient values, was also considered.

2.2.4. TFs based on scalar intensity and gradient magnitude

By combining some results from section 2.2.2 and section 2.2.3, nine different final visualizations by VR were obtained taking into account both information (scalar intensity and gradient magnitude).

2.3. Results

For all the twenty cases, it was observed that $a1$ and $b1$ measured values were at the same position in relation to the respective histogram. Based on this sample, linear regressions were modelled to study the relationship between $a1$ and $b1$ and some statistics (mean, maximum and skewness) of each intensity histogram. Results are presented in Figure 2.2.

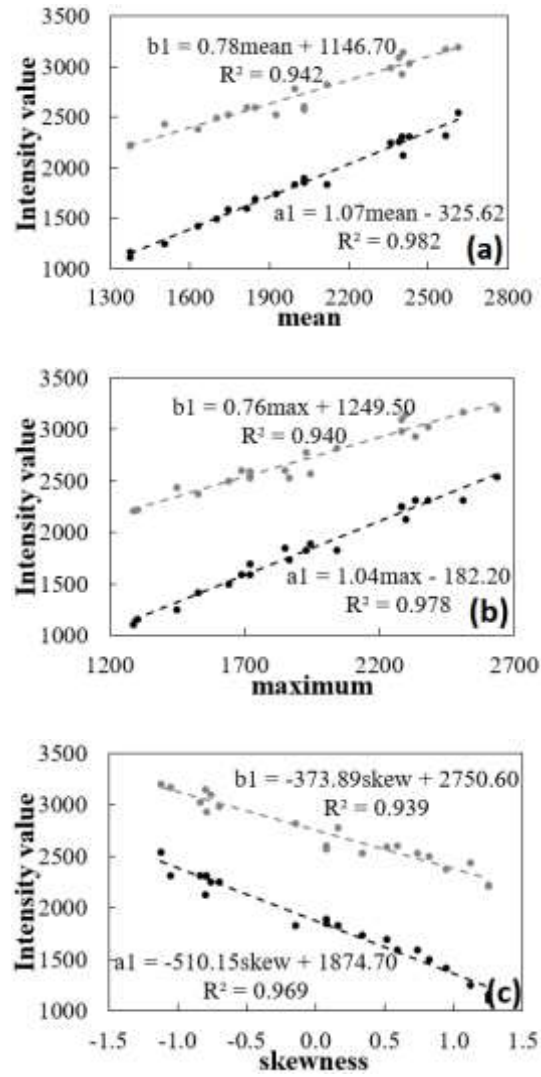


Figure 2.2. Linear regressions obtained between $a1$ and $b1$ and (a) mean, (b) maximum and (c) skewness of histogram intensity values, calculated for each clinical case. Coefficients of determination (R^2) and equations for the estimation of $a1$ and $b1$ based on each statistic are also shown.

The eight combinations of opacity and color TFs based on scalar intensity values ($a1$, $a2$, $b1$ and $b2$) and respective 2D displays of VR results at 0° are shown in Figure 2.3. On the other hand, the VR results obtained with the ten gradient-based TFs, keeping the intensity-based TFs as neutral, are presented in Figure 2.4.

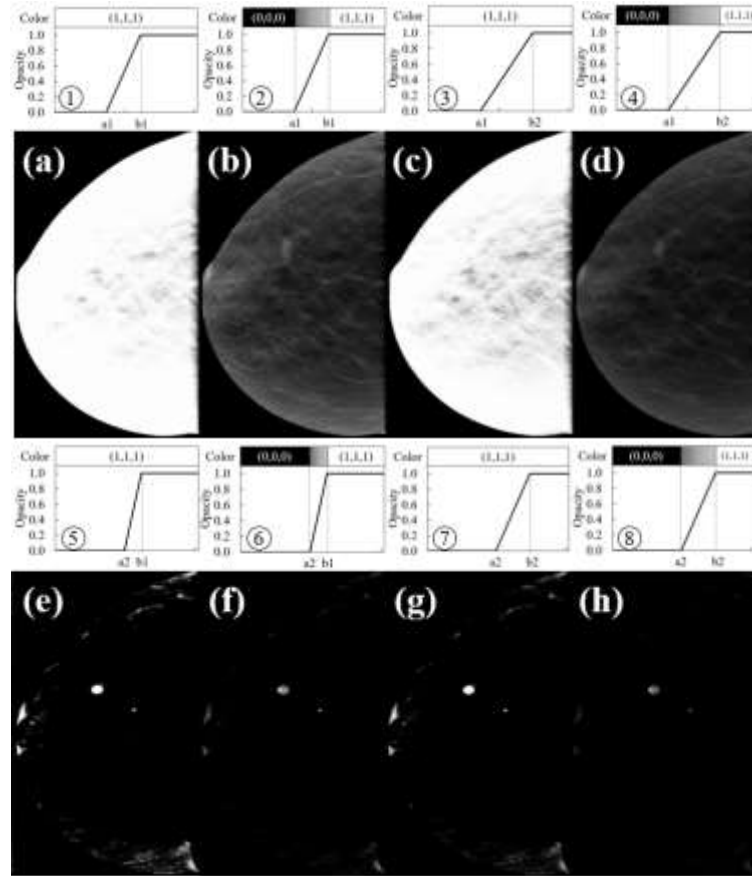
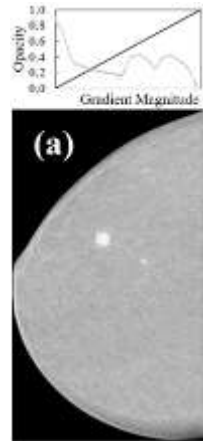


Figure 2.3. 2D displays of VR visualizations at 0° obtained with each combination of intensity-based TFs by varying the color and opacity within the values of a1, a2, b1 and b2.



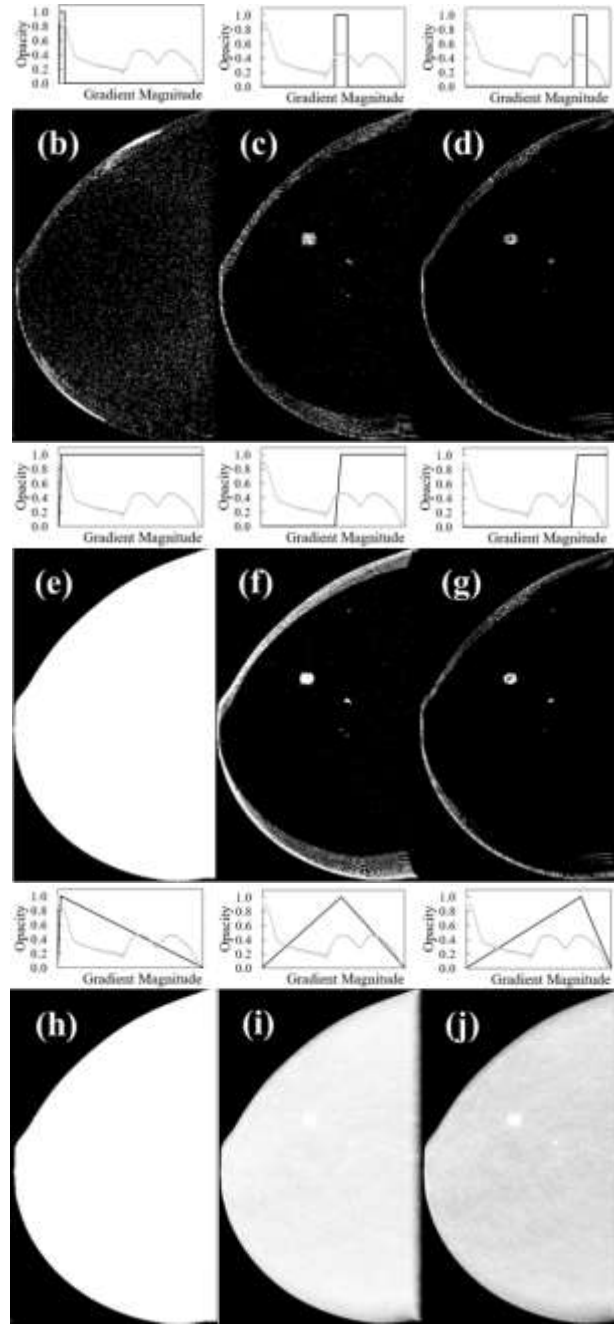


Figure 2.4. 2D displays of VR visualizations at 0° obtained with the ten gradient-based TFs: linear ramp (a), square shaped centered on each peak (b,c,d), ramp-shaped around each peak (e,f,g) and triangular shaped centered on each peak (h,i,j). Opacity and color values were kept at 1 for all scalar intensities.

In order to merge the information provided by the intensity-based functions (Figure 2.3) with the gradient-functions (Figure 2.4), some visualizations of both groups have been selected and combined. From Figure 2.3, it is possible to group three similar visualizations: (1) Figure 2.3 (a) and (c); (2) Figure 2.3 (b) and (d); (3) Figure 2.3 (e-h). From the first group, the combination of functions in Figure 2.3 (a) was selected because it allowed to keep more information about all tissues. From the second one, to have an intermediate visualization, functions of Figure 2.3 (b) were chosen. Of the third group, functions presented in Figure 2.3 (f) were elected because

they emphasize important structures of interest for breast cancer detection, such as calcifications, in a balanced way. Regarding VR resulting from the assignment of opacity to gradient magnitude, there are 7 out of 10 that can be immediately excluded: Figure 2.4 (b-h). TFs corresponding to visualizations of Figure 2.4 (b-d) and Figure 2.4 (f-g) remove a significant amount of information. On the other hand, functions from Figure 2.4 (e) and (h) make too much information opaque (which means that conjugated with intensity-based functions, these gradient information have a minor impact on the final VR). In this way, by combining TFs from Figure 2.3 (a), Figure 2.3 (b) and Figure 2.3 (f) with those of Figure 2.4 (a), Figure 2.4 (i) and Figure 2.4 (j), nine visualizations are obtained as in Figure 2.5.

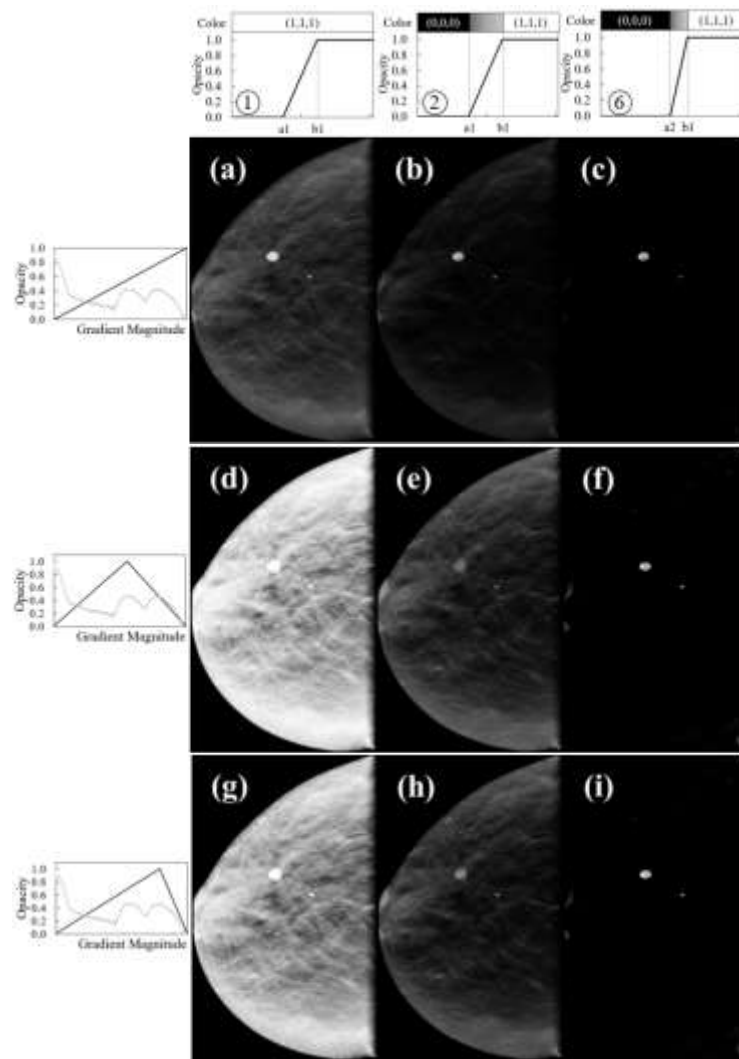


Figure 2.5. 2D displays of VR visualizations at 0° considering intensity and gradient magnitude functions selected. 1st, 2nd and 3rd rows were obtained using the gradient-based functions: linear ramp and triangular-shaped centered on peak 2 and 3, respectively. 1st, 2nd and 3rd columns were obtained using the scalar intensity-based functions: 1, 2 and 6, respectively.

2.4. Discussion and conclusion

A consistent and reproducible methodology for automatic determination of TFs allowing an adequate 3D visualization by VR of DBT data was presented. Opacity/color TFs based on scalar intensity and gradient magnitude were tested separately and together and the results were visualized through composite technique.

Regardless of breast density, all intensity histograms presented the same shape, i.e., neglecting the zero intensity peak corresponding to the black background, a single peak was observed. The measured values for $a1$ and $b1$ were at the same position in relation to the respective histogram. From Figure 2.2, the statistic which presented the best option for automatically estimate $a1$ and $b1$ from the data was the mean, with the highest R^2 simultaneously for the regression with $a1$ and $b1$ (0.9818 and 0.9416, respectively). With the eight VR images observed based on different values of $a1$, $a2$, $b1$ and $b2$ (Figure 2.3), we can see that as the TFs move to the right, the tissues with lower values of intensity become transparent, emphasizing only those of higher intensities, as expected.

As for the intensity, all gradient magnitude histograms presented the same shape for all the twenty cases, independent of breast density. Three main peaks were observed. In a general way, TFs based on gradient magnitude are closely related to the gradient histogram peaks (Figure 2.4). The obtained VR images using only scalar intensity TFs (Figure 2.3 (a), (b) and (f)) can be compared with those using scalar intensity plus gradient magnitude TFs (first, second and third column of Figure 2.5). The general appearance of the latter was determined essentially by TFs based on scalar intensity while the gradient-based TFs improved the definition between regions with different intensities.

Taking into account the preliminary results presented, we conclude that adequate visualization of DBT data by VR can be automatically achieve.

In the future, the impact of VR on detection performance of radiologists and visualization of masses and architectural distortions should be considered through a large and controlled clinical study with 3D VR using DBT data.

Acknowledgments. This work was supported by Universidade de Lisboa (PhD grant) and Fundação para a Ciência e Tecnologia – Portugal (Grant No. SFRH/BD/135733/2018 and FCT-IBEB Strategic Project UID/BIO/00645/2013).

References

1. Gennaro, G., et al., *Digital breast tomosynthesis versus digital mammography: a clinical performance study*. Eur Radiol, 2010. **20**(7): p. 1545-1553.
2. Bonafede, M.M., et al., *Value analysis of digital breast tomosynthesis for breast cancer screening in a commercially-insured US population*. ClinicoEconomics and Outcomes Research: CEOR, 2015. **7**: p. 53-63.
3. Gao, Y., et al., *Digital Breast Tomosynthesis Practice Patterns Following 2011 FDA Approval: A Survey of Breast Imaging Radiologists*. Acad Radiol, 2017. **24**(8): p. 947-953.
4. Destounis, S., A. Santacroce, and A. Arieno, *DBT as a Screening Tool and a Diagnostic Tool*. Current Breast Cancer Reports, 2017. **9**(4): p. 264-271.
5. Samala, R.K., et al., *Digital breast tomosynthesis: Computer-aided detection of clustered microcalcifications on planar projection images*. Phys Med Biol, 2014. **59**(23): p. 7457-7477.
6. Good, W.F., et al., *Digital breast tomosynthesis: a pilot observer study*. AJR Am J Roentgenol, 2008. **190**(4): p. 865-9.
7. Gur, D., et al., *Digital breast tomosynthesis: observer performance study*. AJR Am J Roentgenol, 2009. **193**(2): p. 586-91.
8. Caumo, F., et al., *Digital Breast Tomosynthesis with Synthesized Two-Dimensional Images versus Full-Field Digital Mammography for Population Screening: Outcomes from the Verona Screening Program*. Radiology, 2018. **287**(1): p. 37-46.
9. Venson, J.E., et al., *A Case-Based Study with Radiologists Performing Diagnosis Tasks in Virtual Reality*. Stud Health Technol Inform., 2017. **245**: p. 244-248.
10. Suetens, P., *Medical image analysis*, in *Fundamentals of Medical Imaging* 2009, Cambridge University Press: New York. p. 159-189.
11. Kindlmann, G. and J.W. Durkin. *Semi-automatic generation of transfer functions for direct volume rendering*. in *IEEE Symposium on Volume Visualization (Cat. No.989EX300)*. 1998.
12. Kindlmann, G. *Transfer functions in direct volume rendering: Design, Interface, Interaction*. September 2017 November 2019]; Available from: <http://www.cs.utah.edu/~gk/papers/sig02-TF-notes.pdf>.
13. Kniss, J.K., G.; Hansen, C., *Visualization Handbook*, 2004, Academic Press. p. 181-202.
14. Correa, C.D. and K.L. Ma, *Visibility Histograms and Visibility-Driven Transfer Functions*. IEEE Transactions on Visualization and Computer Graphics, 2011. **17**(2): p. 192-204.
15. Ma, B. and A. Entezari, *Volumetric Feature-Based Classification and Visibility Analysis for Transfer Function Design*. IEEE Transactions on Visualization and Computer Graphics, 2017. **PP**(99): p. 1-1.
16. Alyassin, A.M. *Automatic transfer function generation for volume rendering of high-resolution x-ray 3D digital mammography images*. in *Medical Imaging 2002*. 2002. SPIE.
17. Alyassin, A.M., et al., *3D Visualization of X-ray Tomosynthesis Digital Mammography Data: Preference Study*, in *Digital Mammography: IWDM 2002 — 6th International Workshop on Digital Mammography*, H.-O. Peitgen, Editor 2003, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 507-509.
18. Dharanija, R. and T. Rajalakshmi. *A Conjunct Analysis for Breast Cancer Detection by Volume Rendering of Low Dosage Three Dimensional Mammogram*. in *Progress In Electromagnetics Research Symposium Proceedings*. 2011. China.
19. Jerebko, A., et al. *3D rendering methods for visualization of clusters of calcifications in digital breast tomosynthesis: a feasibility study*. in *ECR 2011*. 2011. Vienna, Austria.
20. Ljung, P., et al., *State of the Art in Transfer Functions for Direct Volume Rendering*. Computer Graphics Forum, 2016. **35**(3): p. 669-691.
21. Schroeder, W., K. Martin, and B. Lorensen, *The Visualization Toolkit: An Object-oriented Approach to 3D Graphics*. 4rd ed 2006, USA: Kitware.
22. VTK - Interpolators. *Visualization Toolkit - VTK - Interpolators*. [cited 2020 February]; Available from: https://vtk.org/Wiki/VTK/Image_Interpolators.

Optimization of Breast Tomosynthesis Visualization through 3D Volume Rendering

3

Ana M. Mota^{1*}, Matthew J. Clarkson², Pedro Almeida¹, and Nuno Matela¹. 2020. "Optimization of Breast Tomosynthesis Visualization through 3D Volume Rendering" *Journal of Imaging* 6, no. 7: 64. <https://doi.org/10.3390/jimaging6070064>

Journal of Imaging is a peer-review journal which is ranked in the **Q2** in “Computer Graphics and Computer-Aided Design”, “Computer Vision and Pattern Recognition” and “Electrical and Electronic Engineering” domains and **Q3** in “Radiology, Nuclear Medicine and Imaging” field.

¹Instituto de Biofísica e Engenharia Biomédica, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal; palmeida@fc.ul.pt (P.A.); nmatela@fc.ul.pt (N.M.)

²Department of Medical Physics and Biomedical Engineering and the Centre for Medical Image Computing, University College London, London WC1E 6BT, UK; m.clarkson@ucl.ac.uk;

* Correspondence: ammota@fc.ul.pt

Abstract

3D volume rendering may represent a complementary option in the visualization of Digital Breast Tomosynthesis (DBT) examinations by providing an understanding of the underlying data at once. Rendering parameters directly influence the quality of rendered images. The purpose of this work is to study the influence of two of these parameters (voxel dimension in z direction and sampling distance) on DBT rendered data. Both parameters were studied with a real phantom and one clinical DBT data set. The voxel size was changed from $0.085 \times 0.085 \times 1.0 \text{ mm}^3$ to $0.085 \times 0.085 \times 0.085 \text{ mm}^3$ using ten interpolation functions available in the Visualization Toolkit library (VTK) and several sampling distance values were evaluated. The results were investigated at 90° using volume rendering visualization with composite technique. For phantom quantitative analysis, degree of smoothness, contrast-to-noise ratio, and full width at half maximum of a Gaussian curve fitted to the profile of one disk were used. Additionally, the time required for each visualization was also recorded. Hamming interpolation function presented the best compromise in image quality. The sampling distance values that showed a better balance between time and image quality were 0.025 mm and 0.05 mm. With the appropriate rendering parameters, a significant improvement in rendered images was achieved.

Keywords: breast tomosynthesis; visualization; volume rendering

3.1. Introduction

Breast cancer remains the most common cancer diagnosed among women and a leading cause of death. However, in the last three decades, there has been a decrease of approximately 40% in the death rate from this disease [1, 2]. This fact is a direct result of the scientific advances in early detection and treatment. Early detection is mostly done through screenings [3, 4]. Until recently, these screenings and breast cancer diagnosis in general were mainly performed by Digital Mammography (DM). However, DM consists of a two-dimensional (2D) acquisition of the three-dimensional (3D) breast causing tissue superposition. This often results in malignant lesions hidden between healthy tissues or normal regions considered as pathological, leading to unnecessary second examinations or biopsies, with additional costs and anxiety for patients [5, 6].

Digital Breast Tomosynthesis (DBT) has consolidated its position as a technique to replace DM in both screening and clinical environments [7-11]. DBT has an acquisition geometry very similar to DM but it acquires a set of projection images, allowing a 3D reconstruction of the breast, reducing the tissue overlap observed with DM [12]. In this way, DBT improves the perception of the location and shape of lesions, without increasing the radiation dose to the patient, when compared to DM [13].

The 3D visualization of DBT is one of the most important and crucial aspects to correctly extract the information provided by this technique. Currently, DBT images are displayed through a 2D slice-by-slice visualization [14], with the analysis done one slice at a time or sequentially as a continuous cine loop, leading to a time-consuming process. Two other new approaches have emerged for the visualization of DBT data: synthetic mammography based on DBT data [15, 16] and thicker slabs obtained by combining several slices [17, 18]. Although synthetic mammography is very useful for comparison with previous DM examinations, it still presents the disadvantages of a 2D visualization (tissue overlapping) [19]. On the other hand, the thicker slabs have revealed good results in reducing time and false positives but have a lower sensitivity [18]. Additionally, a system very recently approved by the Food and Drug Administration, uses artificial intelligence to reduce the number of slices that need to be analyzed by the radiologist. It is based on thicker 6 mm slices combined with synthesized 2D images information [20, 21].

A different type of visualization may play an important complementary role in breast cancer diagnosis [22]. 3D volume rendering is the process of creating realistic computer-generated images of volumetric data, yielding a true depth perception [23]. This type of supplementary visualization is already used in other tomographic medical imaging modalities such as Computed Tomography, Magnetic Resonance Imaging, or Positron Emission Tomography [24-26]. There are some works mentioning a few aspects of 3D volume rendering for DBT [27-29] and its importance to detect clusters of microcalcifications [30].

The classic volume rendering algorithm is ray casting, where rays are cast from the eye or other viewpoints and traverse a scene containing a volumetric data set. This process encompasses several parameters that decide the appearance of the final rendered image. One of these parameters is the sampling distance, which corresponds to the distance between neighboring samples taken along the ray. The value of the sampling distance should be studied and carefully selected accordingly to data set grid resolution [31]. If the distance is too large, our sampling might miss important features in the data and generate major aliasing artifacts. Yet, if we select a very small distance (the number of samples collected along the ray is increased), the amount of time required to render the image will significantly increase [32, 33].

On the other hand, reconstructed DBT data typically has voxel sizes of $0.085 \times 0.085 \times 1.0$ mm³. The anisotropic nature of the reconstructed DBT data is also responsible for serious quality problems in visualization techniques, namely in the direction orthogonal to the detector plane (z-direction). In this way, one solution is to make the grid isotropic through suitable interpolation functions before the rendering process [31]. The smaller the voxel size, the higher the image definition. However, there are more voxels compounding the data set and therefore the processing time of each data set is longer. The time issue is very important in medical image analysis because a large amount of data needs to be displayed and analyzed in real time. For this study, we consider the total time allocated to the visualization process as the sum of interpolation time with rendering time.

In this paper, two parameters that directly affect the quality of the rendered image—sampling distance and reconstructed voxel size—were considered. The main objective was to improve the quality of the rendered images in the z direction and to determine which options allow a better balance between quality and time. In order to transform data to an isotropic grid, several interpolation functions and their corresponding parameters were tested. Additionally, different sampling distance values were introduced in the rendering process. Qualitative and quantitative

analyses of the results were done through visualization by volume rendering of real DBT images of a phantom. Finally, some values were selected and the rendered images obtained for one clinical DBT data set were also analyzed. To the best of our knowledge, this is the first study about optimization of rendering parameters in visualization of DBT data.

3.2. Materials and methods

3.2.1. Data Acquisition and Reconstruction

A phantom made by us was used (Figure 3.1). This phantom consists of an acrylic background to mimic breast tissue, with two embedded columns of aluminum disks to simulate high density lesions (with different diameters and 1.0 mm thick). For this study, the first column of the phantom (Figure 3.1 (b)) was considered.

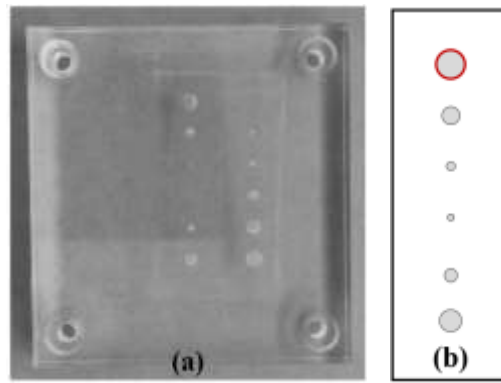


Figure 3.1. (a) Acrylic phantom simulating breast tissue and high density lesions (aluminum disks of different diameters and 1 mm thickness). (b) Scheme of the disks in the first column (top to bottom): 5.0 mm, 3.0 mm, 1.0 mm, 0.5 mm, 2.0 mm, and 4.0 mm, respectively.

Additionally, one clinical DBT data set from an anonymous patient was selected from the clinical facility (Hospital da Luz, Luz Saúde S.A, Lisbon, Portugal) database. Both phantom and clinical data set were acquired with a Siemens MAMMOMAT Inspiration system (Siemens AG, Healthcare Sector, Erlangen, Germany) and reconstructed with the manufacturer algorithm, which uses Filtered Back Projection [34]. The reconstructions have voxel sizes of $0.085 \times 0.085 \times 1.0 \text{ mm}^3$.

3.2.2. Data Visualization

In volume rendering, changing the azimuth of a camera rotates its position around the focal point [32]. In this way, it is possible to have an immediate notion of the entire data volume,

according to several angles. The methods under study in this work have a particular effect on the z -direction. Therefore, we were particularly interested in the visualization perpendicular to the detector plate (with the projection made on the xz planes, along y), i.e., with the camera at 90° (see Figure 3.2). In addition, the visualization with the camera at 0° , that is, parallel to the detector plate (projection is made on the xy planes, along z) was also considered.

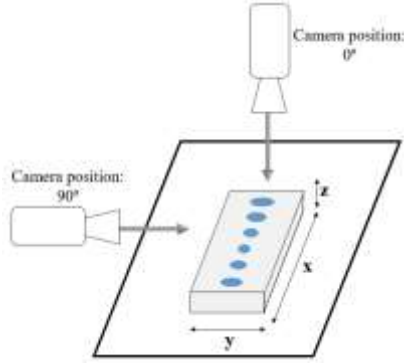


Figure 3.2. Illustrative scheme of visualization at 0° and 90° .

The Visualization Toolkit library (VTK) version 7.1.0. (Kitware, New York, EUA) [32, 35] was used to develop 3D specific software in order to visualize DBT data through volume rendering. The methodologies in study were analyzed using 3D volume rendering visualization with composite technique. An Intel® Core™ i5-5200U CPU (2.20 GHz) @ 8 GB of memory computer was used.

3.2.3. Image Analysis

For phantom quantitative analysis, the profile of the 5.0 mm disk (Figure 3.1 (b) in red) was obtained and three figures of merit were used: full width at half maximum (FWHM) of a Gaussian curve fitted to the 5.0 mm disk's profile, contrast to noise ratio (CNR), and a measure of profile smoothness. $FWHM_{90^\circ}$ was considered as an indicator of the disk's definition at 90° . As already mentioned, DBT presents a lower quality in z than in xy and this is exposed in the spreading of structures of high intensity in z (such as these disks). Although the disks have a thickness of 1.0 mm, the observed values will be higher. In this way, a lower value of $FWHM_{90^\circ}$ will mean a greater definition of the feature at 90° . For CNR, a region of interest (ROI) over the 5.0 mm disk and other two ROIs over the surrounding background were drawn. CNR was calculated using Eq. 3.1:

$$CNR = \frac{\mu_{5.0\text{ mm}} - \mu_{BG}}{\sigma_{BG}} \quad \text{Eq. 3.1}$$

$\mu_{5.0mm}$ and μ_{BG} stand for mean pixel values in ROI over the 5.0 mm disk and background, respectively; and σ_{BG} stands for the mean of standard deviations in background ROIs. In order to obtain a measure for the profile smoothness (important for the interpolation quality analysis), the STEYX Microsoft Excel[®] (Microsoft Office 2013) function was used [36] and its inverse was calculated (Eq. 3.2). The STEYX function gives a measure of the variability of the data in a given range. The degree of smoothness was calculated considering intensity levels between z-distance (16, 24) mm since it corresponded to an area with high variation in intensity:

$$Smoothness = \frac{1}{STEYX_{[16,24]mm}} \quad Eq. 3.2$$

The quantitative analysis was performed at 90°, where the methods under study have the greatest effect. For qualitative purposes, displays of the phantom and one clinical data set obtained with volume rendering visualization at 0° and 90° are presented.

3.2.4. Study of Interpolation Functions

Due to the acquisition process, DBT data has a finer resolution within slices (in *xy* planes) and a coarser resolution between slices (in the *z*-direction). This leads to an anisotropic grid, greatly reducing the quality of rendering. To deal with the anisotropic grid spacing problem, one hypothesis is to change the size of each voxel, considering the smallest dimension, so it can correspond to a perfect cube [37]. Therefore, in this study, the *z*-resolution has been modified to match the resolution within the slice (which has not changed). This results in a homogeneous resolution, improving data quality after reconstruction and before rendering. At the same time, the number of voxels compounding the data set is also increased and thus more computational memory is required. We can refer to a *z*-interpolation, since the voxel size went from $0.085 \times 0.085 \times 1.0 \text{ mm}^3$ to $0.085 \times 0.085 \times 0.085 \text{ mm}^3$, that is, the change occurred only in the third dimension.

As this work is based on C++ software developed with the VTK library, the volume data was resampled into an isotropic grid using appropriate interpolation functions available in this library [38]. In VTK, data interpolation is done internally by several classes. The two main classes considered here were *vtkImageInterpolator* and *vtkImageSincInterpolator*. The first one is the default interpolator and provides linear, cubic, and nearest-neighbor interpolation. The second is responsible for an approximation to sinc interpolation by multiplying one of the available window functions, in order to limit the kernel size. The window functions studied in

this work were: Lanczos, Kaiser (adjustment parameter off), Cosine, Hann, Hamming, Blackman, and Nuttall [38, 39]. In Fourier space, the resolution of spectral window is reduced to the order of the half-width of the sinc function [40]. In this way, the quality of sinc interpolation is related not only to the window function, but also to the window width. For this reason, the window half-width (WHW) was also addressed in this study with values ranging between 1 and 16. *vtkImageSincInterpolator* class has an option to set blur factors in x , y , and z directions in order to blur the data while interpolating [38]. The size of the kernel is automatically increased by the blur factor (BF), increasing the computation time. As here the interpolation is in the z -direction, some values for BF in z were tested. A summary of the functions and parameters analyzed during the interpolation process is shown in Table 3.1.

For the interpolation study, the default value of sampling distance (1.0 mm) was used.

Table 3.1. Summary of the functions and parameters, available in VTK, analyzed during the interpolation process.

In Study	
Image Interpolators	Linear Cubic Nearest-neighbor
Image sinc interpolators	Window Function (Lanczos, Kaiser, Cosine, Hann, Hamming, Blackman, Nuttall) Window Half-Width (WHW) Blur Factor in z -direction (BF(z))

3.2.5. Study of Sampling Distance

In the ray casting volume rendering algorithm, a ray traverses a volume data set. Along the ray, contributions (samples) based on the intensity values weighted by transparency or opacity (transfer functions) are accumulated at discrete locations of the ray, separated by a certain distance – sampling distance. The process of selecting these locations and, therefore, the distance between them, is subject to the sampling theorem. In this case, it is translated on the condition that the distance between two accumulations must be less than or equal to twice the respective smallest voxel spacing [31].

Taking into account the best results obtained with the interpolation functions and their parameters, several sampling distance values were studied. Considering the sampling theorem, values smaller than 2×0.085 mm (0.170 mm) were tested. As previously mentioned, the lower the sampling distance value, the higher the rendering quality should be. As the computation time increases exponentially for lower sampling distance values, our aim was to test several values smaller than 0.170 mm and try to understand where the best balance between quality and

time could be achieved. In addition, five other values outside this range were tested. The default value for sampling distance in VTK is 1.0 mm and it was the maximum value included in our study. Included in the class corresponding to volume ray casting, VTK has an option which automatically computes the sampling distance from the data spacing. Using this option, an automatic sampling distance of 0.195 mm was generated. This value was also included in this work (from a brief previous study, it was concluded that, for our DBT data, this automatically generated value was related to the number of voxels according to an approximation of the expression $\text{autoSD} = 3580.5 \times (\text{nr voxels})^{(-0.621)}$). On the other hand, in order to contextualize the results between 0.195 mm and 1.0 mm, some intermediate values (0.40 mm, 0.60 mm, and 0.80 mm) were also considered. In summary, the sampling distance values analyzed were: 0.010 mm, 0.025 mm, 0.050 mm, 0.075 mm, 0.100 mm, 0.145 mm, 0.170 mm (2×0.085 mm), 0.195 mm (automatic sampling distance), 0.4 mm, 0.6 mm, 0.8 mm, and 1.0 mm (default value in VTK).

3.3. Results

As mentioned, the time involved to render data is very important in order to make visualization through volume rendering useful. Here, we have separated total time spent in visualization as interpolation time plus render time (from the moment the original data is opened until it reaches the screen). The first depends on voxel size and interpolation functions used in the rescaling. The second is related with ray casting process, namely the sampling distance value. Render time values recorded for rendering the original data, as well as data after interpolation (with sampling distance 1.0 mm) are shown in Figure 3.3.

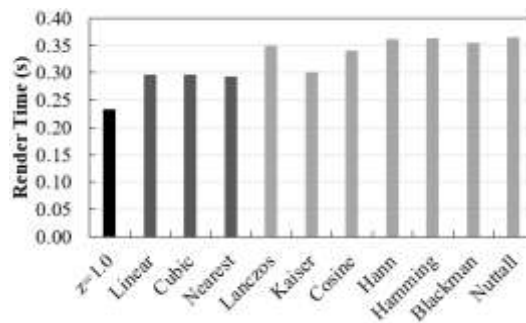


Figure 3.3. Computation time required for rendering the original data (z: 1.0 mm) and data after rescaling with linear, cubic, nearest interpolators and Lanczos, Kaiser, Cosine, Hann, Hamming, Blackman, and Nuttall window functions (with sampling distance 1.0 mm).

As the interpolation was done only in the third dimension, all the results presented in this section were measured considering phantom visualization at 90° (i.e., through the z direction).

3.3.1. Study of Image Interpolators

3.3.1.1. Linear, Cubic and Nearest-Neighbor Interpolation

The 5.0 mm disk's profiles obtained after rescaling with the linear, cubic, and nearest-neighbor interpolators are presented in Figure 3.4. Additionally, it is also shown the profile obtained before rescaling (with voxel size of $0.085 \times 0.085 \times 1.0 \text{ mm}^3$). FWHM and smoothness values measured for each visualization at 90° are presented in Table 3.2. The total time required in the process is also shown and it stands for the rendered time (Figure 3.3) plus interpolation time.

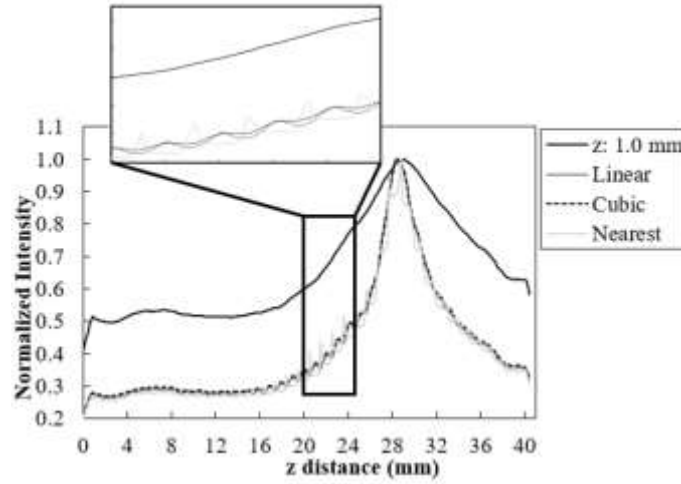


Figure 3.4. Profile of the 5.0 mm disk obtained at 90° for the original data (z : 1.0 mm) and after rescaling with the linear, cubic, and nearest interpolators. Zoom-in of a range with large intensity variation.

Table 3.2. FWHM and smoothness values measured in interpolated data at 90° . The total time (interpolation plus rendering time) is also presented.

Interpolator	Linear	Cubic	Nearest
Total time (secs)	0.45	0.50	0.45
FWHM $_{90^\circ}$ (mm)	7.55	7.79	8.44
Smoothness $_{90^\circ}$	64.5	72.2	37.4

3.3.1.2. Sinc Interpolation with Different Window Functions

The sinc function was multiplied by different window functions (Lanczos, Kaiser, Cosine, Hann, Hamming, Blackman, and Nuttall). The 5.0 mm disk's profiles obtained with WHW values ranging from 1 to 16 are presented in Figure 3.5.

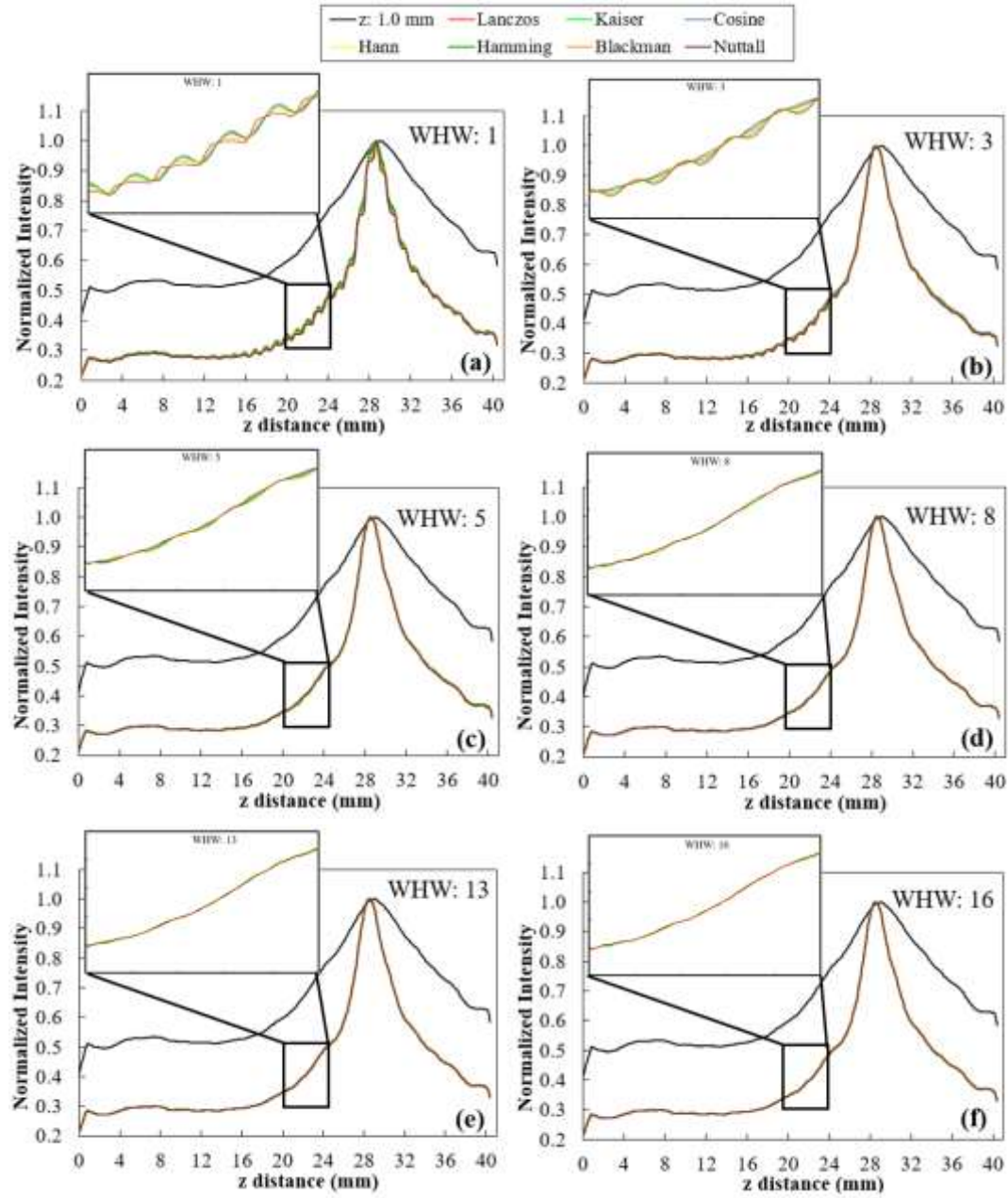


Figure 3.5. Profiles of the 5.0 mm disk obtained at 90° for the original data (z : 1.0 mm) and after rescaling with Lanczos, Kaiser, Cosine, Hann, Hamming, Blackman, and Nuttall window functions with WHW values of 1, 3, 5, 8, 13, and 16 ((a) to (f), respectively). Zoom-in of a range with large intensity variation.

Based on profiles of Figure 3.5 and for each window function and WHW, the corresponding smoothness and FWHM values were determined and the results are displayed in Figure 3.6.

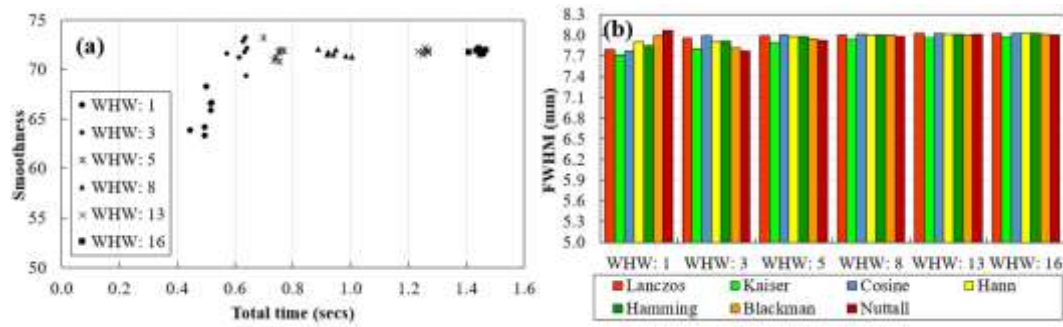
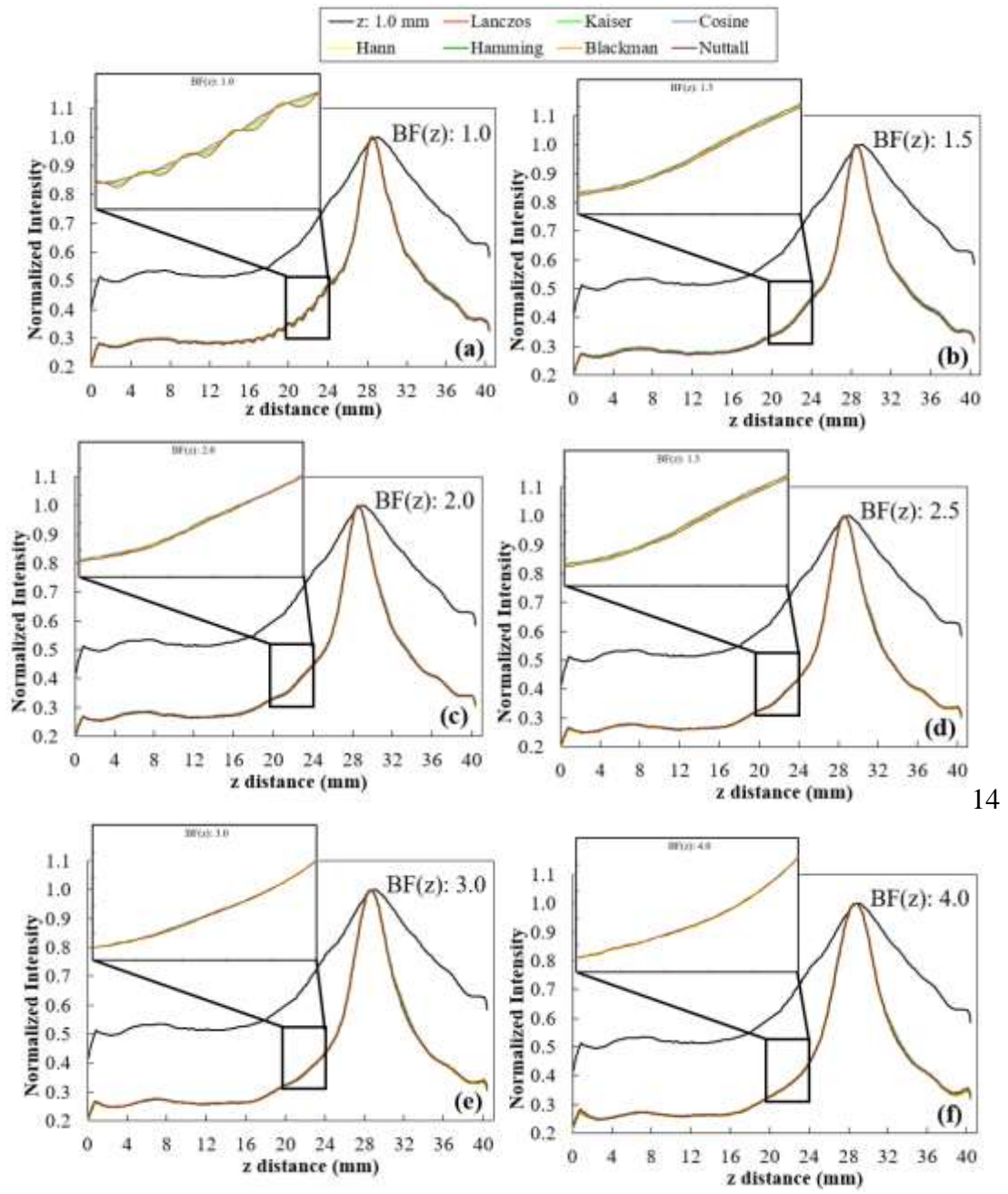


Figure 3.6. Results obtained with Lanczos, Kaiser, Cosine, Hann, Hamming, Blackman, and Nuttall window functions for the profile at 90° , by changing WHW values (from 1 to 16). (a) Smoothness values as a function of total time and (b) FWHM of the 5.0 mm disk at 90° .

The profiles calculated with BF (z) values ranging from 1.0 to 4.0 are shown in Figure 3.7 from (a) to (f), respectively. As for the WHW, based on profiles of Figure 3.7, the corresponding smoothness and FWHM values for each BF (z) and window function were determined and the results are presented in Figure 3.8.



14

Figure 3.7. Profiles of the 5-mm disk obtained at 90° for the original data (z: 1.0 mm) and after rescaling with Lanczos, Kaiser, Cosine, Hann, Hamming, Blackman, and Nuttall window functions with BF(z) values of 1.0, 1.5, 2.0, 2.5, 3.0, and 4.0 ((a)–(f)). Zoom-in of a range with large intensity variation.

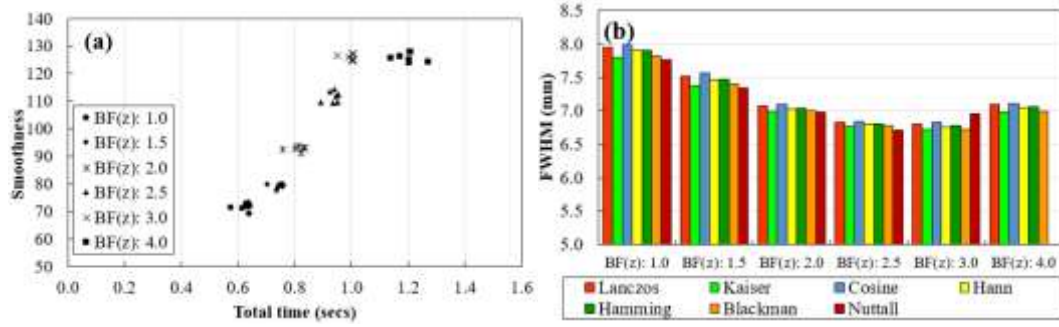


Figure 3.8. Results obtained with Lanczos, Kaiser, Cosine, Hann, Hamming, Blackman, and Nuttall window functions for the profile at 90° , by changing BF(z) values (from 1.0 to 4.0). (a) Smoothness values as a function of total time and (b) FWHM of the 5.0 mm disk at 90° .

3.3.2. Selection of Functions and Parameters for the Sampling Distance Study

The study of interpolation functions and their parameters is, by itself, quite extensive. Thus, to simplify the analysis of the results obtained with different sampling distance values, some selections were made to proceed.

- **WHW:** In Figure 3.5, for WHW values above 5, the obtained profiles are very similar. This is translated into the results of Figure 3.6 (a), where $WHW = 5$ produces the highest smoothness in the shortest time. From this value on, there is no significant increase in smoothness, only an increase in the time required for interpolation. On the other hand, in Figure 3.6 (b), the FWHM values are very similar for the different WHW values. In this way, the choice is based on the smoothest profile in the shortest possible time. This corresponds to $WHW = 5$.
- **BF(z):** According to Figure 3.7, for $BF(z) \geq 1.5$, there is a significant decrease in the variability of the profiles. Through the results in Figure 3.8 (a), it is possible to observe that the higher the value of BF (z), the greater the smoothness, reaching a certain convergence for $BF(z) \geq 3$. From $BF(z) = 1.0$ to $BF(z) = 2.0$, there is a visible decrease in the value of FWHM (Figure 3.8 (b)), and for $BF(z) > 2$, these values become very similar. Thus, we choose $BF(z) = 2$ as the value that represents a better compromise between smoothness, FWHM, and time.
- **Interpolator:** Nearest and linear interpolators were excluded since the corresponding profiles showed low smoothness when compared to the others. The cubic interpolator was selected to proceed as it presented smoothness and FWHM values comparable to the other interpolations with a similar interpolation time. For the sinc interpolator,

considering the results with $WHW = 5$ and $BF(z) = 2$, window functions were sorted according to the profile smoothness (in decreasing order). The Hamming window function presented the best correspondent result between the two options (for $WHW = 5$: Kaiser > Nuttall > Hamming; for $BF(z) = 2$: Blackman > Lanczos > Hann > Hamming). Since, among them, window functions presented very close results, this selection of a single function to proceed was done only to simplify and concentrate the next results.

In summary, the next section results are for cubic interpolation and Hamming window function (sinc) with $WHW = 5$ or/and $BF(z) = 2$.

3.3.3. Sampling Distance Study

For each sampling distance, the time required to render the original data (z : 1.0 mm, black in Figure 3.9) was recorded. In addition, the average rendering times of each selected interpolator (cubic, Hamming with $WHW = 5$, Hamming with $BF(z) = 2$ and Hamming with $WHW = 5$ and $BF(z) = 2$) were also calculated (gray in Figure 3.9).

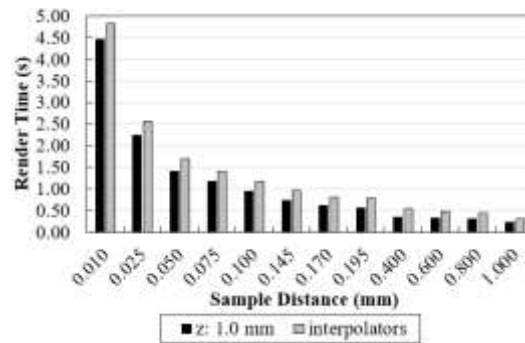


Figure 3.9. Computation time required for rendering the original data (black) and data after rescaling (gray) taking into account the different sampling distance values. For each sampling distance, each gray value was obtained by averaging the rendering times recorded for each interpolator considered here (cubic, Hamming with $WHW = 5$, Hamming with $BF(z) = 2$ and Hamming with $WHW = 5$ and $BF(z) = 2$).

The 5.0 mm disk's profiles were obtained from the rendered images with each sampling distance value and for each interpolator. The measured profiles between 20 mm and 36 mm (z distance) are shown in Figure 3.10.

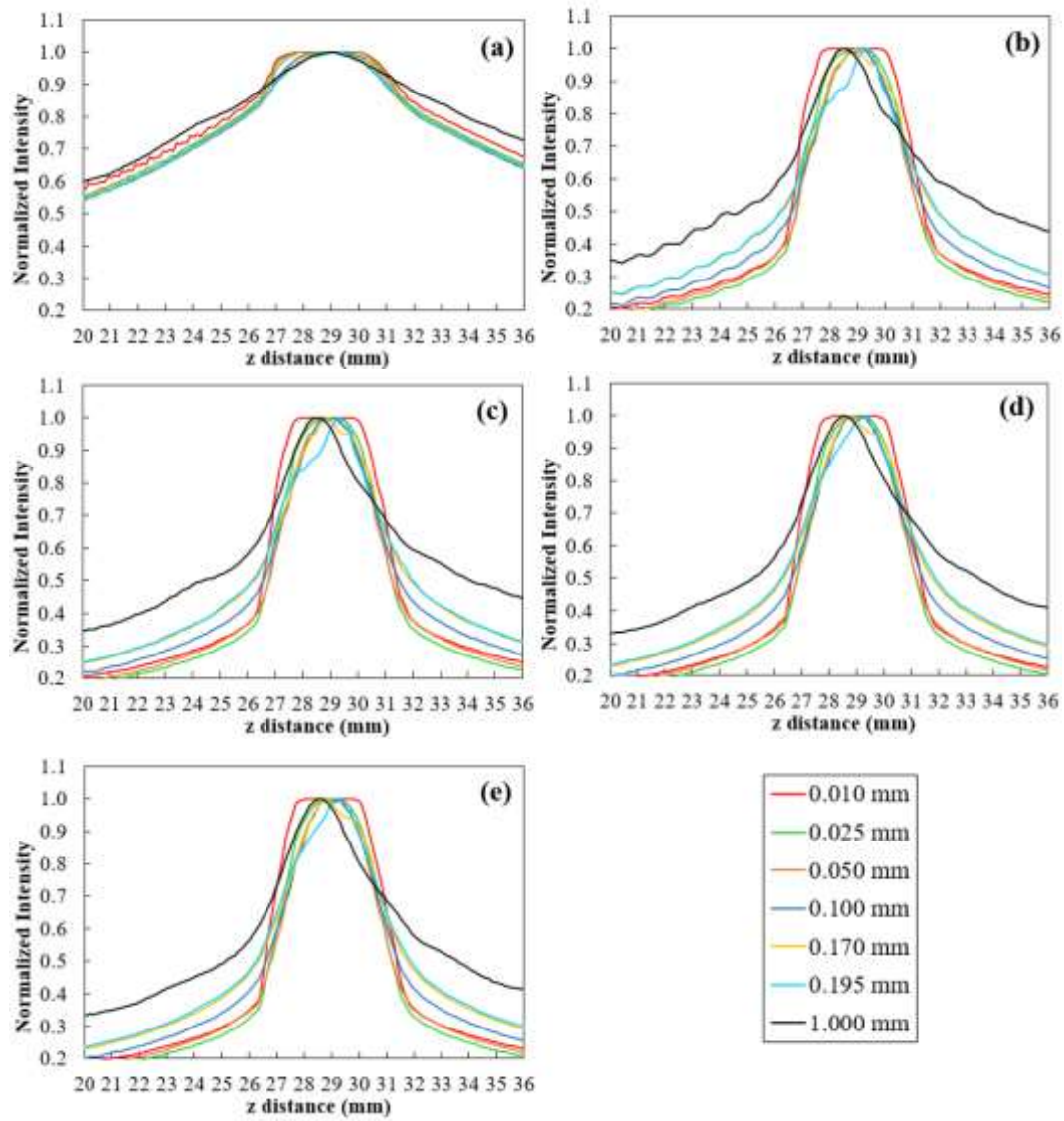


Figure 3.10. Profiles of the 5.0 mm disk obtained at 90° , for some values of sampling distance tested (0.010, 0.025, 0.050, 0.100, 0.170, 0.195, and 1.0 mm), with the original data (a) and after rescaling with cubic (b), Hamming with WHW = 5 (c), Hamming with BF (z) = 2 (d), and Hamming with WHW = 5 and BF (z) = 2 (e).

For a quantitative analysis of the quality of rendered images, smoothness, CNR, and FWHM values were calculated for each sampling distance and the results obtained with the original data and after interpolation are shown in Figure 3.11.

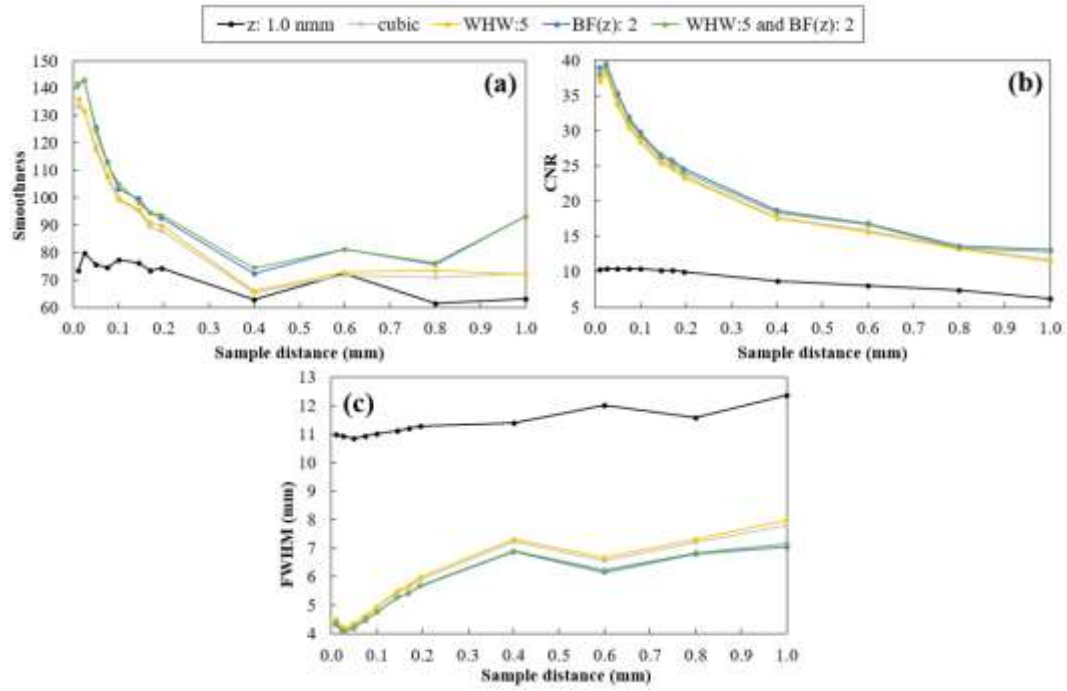


Figure 3.11. Smoothness (a), CNR (b) and FWHM (c) plotted as a function of sampling distance values for original data and data after interpolation with cubic, Hamming with WHW = 5, Hamming with BF (z) = 2 and Hamming with WHW = 5 and BF (z) = 2. Results obtained for rendered images at 90° .

For qualitative inspection, sinc interpolation with Hamming window and BF (z) = 2 were selected and sampling distance of 0.025 mm were used. Images of the 5.0 mm disk of the phantom are presented in Figure 3.12. Images achieved with volume rendering at 0° and 90° are shown in the first and second column of Figure 3.12, respectively. The first row represents original data with default visualization (without interpolation and with sampling distance 1.0 mm), the second row presents “processed” data with Hamming BF (z) = 2 interpolation, and sampling distance 0.025 mm. The quantitative analysis corresponding to the images in Figure 3.12 is summarized in Table 3.3.

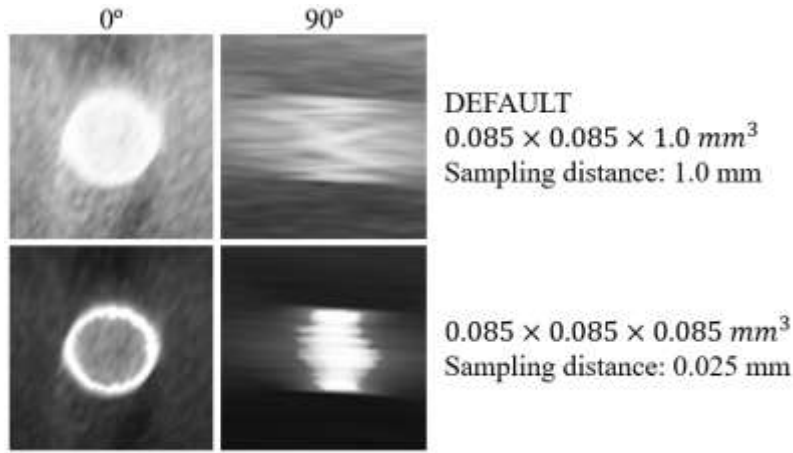


Figure 3.12. Volume rendering images at 0° and 90° for the 5-mm disk obtained for the original data with default visualization (top row) and interpolated data with Hamming window with BF (z) = 2 and sampling distance 0.025 mm (bottom row).

Table 3.3. Summary of the results obtained for phantom at 0° and 90°, with default visualization options (voxels with $0.085 \times 0.085 \times 1.0 \text{ mm}^3$ and sampling distance 1.0 mm) and the options selected in our study (voxels with $0.085 \times 0.085 \times 0.085 \text{ mm}^3$ after interpolation with Hamming window function and BF (z) = 2 and sampling distance 0.025 mm).

	Default	From Our Study
Voxel size (mm^3)	$0.085 \times 0.085 \times 1.0$	$0.085 \times 0.085 \times 0.085$ (Hamming with BF (z) = 2)
Sampling distance (mm)	1.0	0.025
Total time (s)	0.23	3.05
CNR_{0°	7.19	22.12
FWHM_{0° (mm)	3.67	3.52
CNR_{90°	6.23	39.39
FWHM_{90° (mm)	12.38	4.06
Smoothness $_{90^\circ}$	63.0	142.8

3.3.4. Clinical Data

To evaluate the consistency of the results, sinc interpolator with Hamming window function with BF (z) = 2 was used in the rescaling of one clinical case ($0.085 \times 0.085 \times 1.0 \text{ mm}^3 \rightarrow 0.085 \times 0.085 \times 0.085 \text{ mm}^3$). Volume rendering of clinical data was obtained with sampling distance value of 0.025 mm since it showed good results in terms of smoothness, CNR, and FWHM. 2D displays of composite volume rendering of clinical data obtained at 0° and 90° are shown in Figure 3.13.

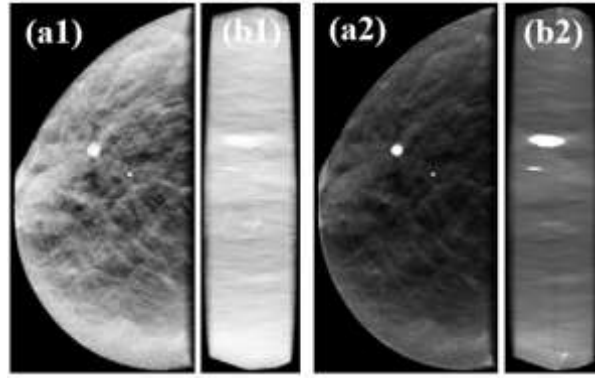


Figure 3.13. 2D displays of composite volume rendering visualization obtained at 0° and 90° (a and b, respectively) for original data with default sampling distance (1.0 mm) (a1 and b1) and interpolated data with sampling distance 0.025 mm (a2 and b2).

3.4. Discussion

This type of visualization is an alternative and complementary approach to the standard time consuming slice-by-slice visualization. Here, we optimized the volume rendering visualization for DBT data and our analysis was focused on two main parameters: interpolation methods used before rendering to obtain an isotropic grid (by modifying reconstructed voxel size); and sampling distance values.

For a constant sampling distance in the interpolation study, rendering time proved to be similar between the different interpolation methods (the minimum value recorded was 0.29 seconds for nearest and the maximum was 0.37 s for Nuttall) (Figure 3.3). On the other hand, when the number of voxels increases from approximately 8 million ($0.085 \times 0.085 \times 1.0 \text{ mm}^3$) to 95 million ($0.085 \times 0.085 \times 0.085 \text{ mm}^3$), a greater difference was observed. This means that changing the number of voxels had a greater impact on rendering time than the interpolation method itself.

Taking into account the profiles in Figure 3.4, with a smaller voxel size, a narrower profile was observed, confirming an improvement of image resolution. Among the three interpolators from *vtkImageInterpolator* class, nearest presented a profile with a blocky appearance resulting from the discontinuous interpolation between neighboring voxels. This translates into a lower smoothness value (greater variability) and a higher FWHM value (Table 3.2). Linear and cubic interpolators showed similar results for FWHM, with cubic showing a greater smoothness of the profile.

Window functions considered here (Lanczos, Kaiser, Cosine, Hann, Hamming, Blackman, and Nuttall) are already a selection of functions that produce high quality interpolations. Thus, by expecting a similar quality resultant from the different functions, our analysis was focused on two influential parameters: WHW and BF (z). By default, these options are set to 3 and 1.0, respectively, in VTK. Starting with WHW, there was a visible difference in the variability of the profiles in Figure 3.5, in particular until $\text{WHW} = 5$, with no major differences between interpolators. This observation was corroborated by the calculation of smoothness and FWHM values shown in Figure 3.6. In Figure 3.6 (a), there are six distinct groups, corresponding to the six WHW values tested and, for WHW values above 5, there was no noticeable increase in the smoothness value, increasing only the interpolation time. On the other hand, the variation of this parameter did not cause large fluctuations in the FWHM values for the different interpolators, as can be seen in Figure 3.6 (b). About the BF in z , it was found that this parameter has a great influence on the intensity fluctuations existent in the images. For example, from BF (z) = 1.0 (default) to BF (z) = 1.5 (Figure 3.7 (a) and (b), respectively), there was a significant modification in the profile of the disk at 90° . The results observed in the profiles were confirmed by the numerical analysis of smoothness (Figure 3.8 (a)) which increases significantly for higher BF (z) values. It would be expected that the introduction of a blur parameter would increase the smoothness by compromising dispersion in the z -direction (increasing the FWHM value). However, as seen in Figure 3.8 (b), the introduction of this factor in the interpolation (up to a certain limit) helps delineating the structures during rendering. Until BF (z) = 2.5, there is a marked decrease in FWHM values for all window functions. Taking into account the temporal information of Figure 3.8 (a) on the x -axis, we can see that these improvements were achieved with an increase in the interpolation time.

In Section 3.3.2, the selection considered for the sampling distance study has already been explained ($\text{WHW} = 5$, $\text{BF}(z) = 2.0$ and Hamming window function). While the choice of $\text{WHW} = 5$ was quite simple, for BF (z), we could have opted for 2.0 or 2.5. We selected the first instead of the second value because the interpolation time was shorter. As for the window function in sinc interpolator, it was not an obvious selection, since the different functions presented very similar results. In addition, for example, the Kaiser function has a parameter (α) responsible for the balance between blurring and ringing [31] and so, the effect of α on the results should be considered in a future work. Despite this and in view of the obtained results, Hamming function was the one with the best quality considering both $\text{WHW} = 5$ and $\text{BF}(z) = 2$.

By analyzing the time results obtained for various sampling distance values, if we compare Figure 3.3 and Figure 3.9, we found that sampling distance had much more impact on the rendering time than the number of voxels involved. Figure 3.9 shows an exponential curve with lower sampling distance values corresponding to much longer times. This was expected because sampling distance has a predominant role in the ray casting process.

In Figure 3.10, it is possible to notice once again the difference in the profiles obtained before and after rescaling (as it had already been observed before in Figure 3.5 and Figure 3.7 with sampling distance 1.0 mm). As for the interpolated data, in contrast to the profiles obtained with Hamming, the cubic interpolator showed some oscillation in intensity between 20 mm and 26 mm for all sampling distance values (Figure 3.10 (b)). On the other hand, it is evident that the smaller the sampling distance, the greater the definition of the plateau, which corresponds to a well-defined disk in the rendered image. In Figure 3.11, the highlight goes to sampling distance of 0.025 mm, which shows a peak in smoothness and CNR and the lowest value of FWHM. In the same figure, taking into account the different sampling distances, the results for the four interpolations showed to be similar for CNR and FWHM. For the smoothness level of the profiles, Hamming with $BF(z) = 2$ and Hamming with $BF(z) = 2$ and $WHW = 5$ stood out, with the first one presenting the best results in general. All four interpolations lead to a significant improvement in the quality, when compared to the original data (black in Figure 3.11). This improvement is visible in the rendered images of Figure 3.12 where, particularly at 90° , the thickness of the disk can be observed with great definition (FWHM of 4.06 mm -Table 3.3), while with default values what we see is a blurred disk (FWHM of 12.38 mm -Table 3.3). This translates into a decrease of about 67% in FWHM values. Furthermore, there was an increase in CNR and smoothness of around 500% and 127%, respectively. At 0° , there is also an increase in CNR and a slight decrease in FWHM (about 4%). On the other hand, with interpolation and sampling distance of 0.025 mm, the time required for all the visualization process increases by approximately a factor of 10. Despite remaining at an acceptable value, time continues to be a crucial parameter in rendering and should be optimized.

Another viable option for the sampling distance could be 0.05 mm as it is 1.4 times faster than 0.025 mm, despite suffering some losses in image quality. Here is a summary of the numerical results obtained for the image quality with sampling distance of 0.05 mm (for comparison with the last column of Table 3.3): Total time: 2.19 s; CNR_{0° : 22.65; $FWHM_{0^\circ}$: 3.52 mm; CNR_{90° : 35.35; $FWHM_{90^\circ}$: 4.21 mm; and $Smoothness_{90^\circ}$: 125.7. Taking into account these results, for

some applications the balance between image quality and time may compensate with a sampling distance of 0.05 mm.

A sampling distance of 0.025 mm and rescaling with Hamming BF (z) = 2 were applied to the visualization of a clinical data set. Rendered images of the clinical data (Figure 3.13) were in accordance with the results obtained with the phantom in terms of improving image quality, namely at 90°. There was a visual increase in contrast in the final image (Figure 3.13 (b)) and the calcification present in this case had showed better defined contours. The preliminary clinical results are intended to consolidate the results obtained with the phantom. In the future, additional accurate studies with volume rendering of DBT clinical data should be considered to perform quantitative analysis of this type of visualization.

3.5. Conclusions

DBT visualization by volume rendering is a new field of research that may support breast cancer diagnosis. With this type of visualization, there is the advantage of observing the entire volume data set at once, from different angles. This is a complex process of visualizing volumetric data set, which includes several factors crucial for the final rendered image. The optimization of these factors is extremely important so that visualization through volume rendering can, along with 2D visualization, have true clinical value. In this way, the main objective of this work was to study and optimize two of these parameters: the interpolation used in the transformation of an anisotropic into an isotropic grid and appropriate sampling distance values, taking into account the entire time required. This is a very extensive study, so the analysis of other interpolation functions and their parameters should be considered in future work.

Author Contributions: All the authors substantially contributed to this paper. Formal analysis, A.M.M.; Methodology, A.M.M.; Supervision, M.J.C., P.A. and N.M.; Writing—original draft, A.M.M.; Writing—review & editing, N.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by Universidade de Lisboa (PhD grant) and Fundação para a Ciência e Tecnologia–Portugal (Grant No. SFRH/BD/135733/2018 and FCT-IBEB Strategic Project UIDB/00645/2020).

Acknowledgments: The authors would like to thank Luis Peralta, who was responsible for the construction of the phantom used.

References

1. Ferlay, J., et al., *Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018*. European Journal of Cancer, 2018. **103**: p. 356-387.
2. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2020*. CA Cancer J Clin, 2020. **70**(1): p. 7-30.
3. Berry, D.A., et al., *Effect of screening and adjuvant therapy on mortality from breast cancer*. N Engl J Med, 2005. **353**(17): p. 1784-92.
4. Independent UK Panel on Breast Cancer Screening, *The benefits and harms of breast cancer screening: an independent review*. The Lancet, 2012. **380**(9855): p. 1778-1786.
5. Poplack, S.P., et al., *Digital breast tomosynthesis: initial experience in 98 women with abnormal digital screening mammography*. AJR Am J Roentgenol, 2007. **189**(3): p. 616-23.
6. Hubbard, R.A., et al., *Cumulative Probability of False-Positive Recall or Biopsy Recommendation After 10 Years of Screening Mammography A Cohort Study*. Annals of Internal Medicine, 2011. **155**(8): p. 481-492.
7. Gennaro, G., et al., *Digital breast tomosynthesis versus digital mammography: a clinical performance study*. Eur Radiol, 2010. **20**(7): p. 1545-1553.
8. Brandt, K.R., et al., *Can Digital Breast Tomosynthesis Replace Conventional Diagnostic Mammography Views for Screening Recalls Without Calcifications? A Comparison Study in a Simulated Clinical Setting*. American Journal of Roentgenology, 2013. **200**(2): p. 291-298.
9. Bonafede, M.M., et al., *Value analysis of digital breast tomosynthesis for breast cancer screening in a commercially-insured US population*. ClinicoEconomics and Outcomes Research: CEOR, 2015. **7**: p. 53-63.
10. Gao, Y., et al., *Digital Breast Tomosynthesis Practice Patterns Following 2011 FDA Approval: A Survey of Breast Imaging Radiologists*. Acad Radiol, 2017. **24**(8): p. 947-953.
11. Destounis, S., A. Santacroce, and A. Arieno, *DBT as a Screening Tool and a Diagnostic Tool*. Current Breast Cancer Reports, 2017. **9**(4): p. 264-271.
12. Ramasundara, S., et al., *Diagnostic implications of digital breast tomosynthesis in symptomatic patients*. Breast Cancer Research : BCR, 2015. **17**(Suppl 1): p. P20-P20.
13. Svahn, T.M., et al., *Review of radiation dose estimates in digital breast tomosynthesis relative to those in two-view full-field digital mammography*. The Breast, 2015. **24**(2): p. 93-99.
14. Sechopoulos, I., *A review of breast tomosynthesis. Part I. The image acquisition process*. Med Phys, 2013. **40**(1): p. 014301.
15. Hofvind, S., et al., *Digital Breast Tomosynthesis and Synthetic 2D Mammography versus Digital Mammography: Evaluation in a Population-based Screening Program*. Radiology, 2018. **287**(3): p. 787-794.
16. Simon, K., et al., *Accuracy of Synthetic 2D Mammography Compared With Conventional 2D Digital Mammography Obtained With 3D Tomosynthesis*. American Journal of Roentgenology, 2019. **212**(6): p. 1406-1411.
17. van Schie, G., et al., *Mass detection in reconstructed digital breast tomosynthesis volumes with a computer-aided detection system trained on 2D mammograms*. Med Phys, 2013. **40**(4): p. 041902.
18. Iotti, V., et al., *Comparing two visualization protocols for tomosynthesis in screening: specificity and sensitivity of slabs versus planes plus slabs*. Eur Radiol, 2019. **29**(7): p. 3802-3811.
19. Petropoulos, A.E., et al., *Quantitative assessment of microcalcification cluster image quality in digital breast tomosynthesis, 2-dimensional and synthetic mammography*. Medical & Biological Engineering & Computing, 2020. **58**(1): p. 187-209.

20. Food and Drug Administration (FDA) U.S. . *Approval for software option 3DQuorum™ technology - Premarket Approval*. 2019 [cited 2020 25 June]; Available from: <https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfpma/pma.cfm?id=P080003S008>.
21. *3DQuorum™ Imaging Technology - Improving radiologist performance through Artificial Intelligence and SmartSlices (White Paper)*. 2020; Available from: [https://www.hologic.com/sites/default/files/downloads/WP-00152_Rev001_3DQuorum_Imaging_Technology_Whitepaper%20%20\(1\).pdf](https://www.hologic.com/sites/default/files/downloads/WP-00152_Rev001_3DQuorum_Imaging_Technology_Whitepaper%20%20(1).pdf).
22. Venson, J.E., et al., *A Case-Based Study with Radiologists Performing Diagnosis Tasks in Virtual Reality*. *Stud Health Technol Inform.*, 2017. **245**: p. 244-248.
23. Suetens, P., *Medical image analysis, in Fundamentals of Medical Imaging* 2009, Cambridge University Press: New York. p. 159-189.
24. O'Connell, A., et al., *Cone-Beam CT for Breast Imaging: Radiation Dose, Breast Coverage, and Image Quality*. *American Journal of Roentgenology*, 2010. **195**(2): p. 496-509.
25. Song, H., X. Cui, and F. Sun, *Breast Tissue 3D Segmentation and Visualization on MRI*. *International Journal of Biomedical Imaging*, 2013. **2013**: p. 8.
26. Jung, Y., et al., *Occlusion and Slice-Based Volume Rendering Augmentation for PET-CT*. *IEEE Journal of Biomedical and Health Informatics*, 2017. **21**(4): p. 1005-1014.
27. Alyassin, A.M. *Automatic transfer function generation for volume rendering of high-resolution x-ray 3D digital mammography images*. in *Medical Imaging 2002*. 2002. SPIE.
28. Alyassin, A.M., et al., *3D Visualization of X-ray Tomosynthesis Digital Mammography Data: Preference Study*, in *Digital Mammography: IWDM 2002 — 6th International Workshop on Digital Mammography*, H.-O. Peitgen, Editor 2003, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 507-509.
29. Dharanija, R. and T. Rajalakshmi. *A Conjunct Analysis for Breast Cancer Detection by Volume Rendering of Low Dosage Three Dimensional Mammogram*. in *Progress In Electromagnetics Research Symposium Proceedings*. 2011. China.
30. Jerebko, A., et al. *3D rendering methods for visualization of clusters of calcifications in digital breast tomosynthesis: a feasibility study*. in *ECR 2011*. 2011. Vienna, Austria.
31. Preim, B. and D. Bartz, *Visualization in Medicine: Theory, Algorithms, and Applications* 2007: Morgan Kaufmann.
32. Schroeder, W., K. Martin, and B. Lorensen, *The Visualization Toolkit: An Object-oriented Approach to 3D Graphics*. 4rd ed 2006, USA: Kitware.
33. Kitware, *The VTK User's Guide*. 11th ed 2010, USA: Kitware.
34. Siemens. *MAMMOMAT Inspiration - Tomosynthesis Option*. 2015 [cited 2020 February]; Available from: https://www.accessdata.fda.gov/cdrh_docs/pdf14/P140011c.pdf.
35. VTK. *Visualization Toolkit - VTK*. 2020 [cited 2020 February]; Available from: <http://www.vtk.org/>.
36. *STEYX function*. 2020 [cited 2020 April]; Available from: <https://support.microsoft.com/en-us/office/steyx-function-6ce74b2c-449d-4a6e-b9ac-f9cef5ba48ab?ui=en-us&rs=en-us&ad=us>.
37. Thévenaz, P., T. Blu, and M. Unser, *Chapter 28 - Image Interpolation and Resampling*, in *Handbook of Medical Image Processing and Analysis (Second Edition)*, I.N. Bankman, Editor 2009, Academic Press: Burlington. p. 465-493.
38. VTK - Interpolators. *Visualization Toolkit - VTK - Interpolators*. [cited 2020 February]; Available from: https://vtk.org/Wiki/VTK/Image_Interpolators.
39. Nuttall, A., *Some windows with very good sidelobe behavior*. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 1981. **29**(1): p. 84-91.
40. Jähne, B., 2.3.3. *The Sampling Theorem*, in *Digital Image Processing: Concepts, Algorithms, and Scientific Applications* 1995, Springer: Berlin, Germany.

Impact of total variation minimization in volume rendering visualization of breast tomosynthesis data

4

Ana M. Mota^{a*}, Matthew J. Clarkson^b, Pedro Almeida^a, Luis Peralta^{c,d} and Nuno Matela^a. “Impact of total variation minimization in volume rendering visualization of breast tomosynthesis data”. *Computer Methods and Programs in Biomedicine*. 2020 Oct; 195:105534. doi: 10.1016/j.cmpb.2020.105534.

Computer Methods and Programs in Biomedicine is a peer-review journal which have an impact factor of **5.428** and is ranked in the **Q1** in the domains: “Computer science – interdisciplinary applications”, “Computer science – theory & methods”, “Biomedical Engineering” and “Medical Informatics”.

^aFaculdade de Ciências, Instituto de Biofísica e Engenharia Biomédica, Universidade de Lisboa, Campo Grande 1749-016 Lisboa, Portugal;

^bDepartment of Medical Physics and Biomedical Engineering and the Centre for Medical Image Computing, University College London, London, UK;

^cDepartamento de Física da Faculdade de Ciências da Universidade de Lisboa, Lisboa, Portugal

^dLaboratório de Instrumentação e Física Experimental de Partículas, Lisboa, Portugal

* Corresponding author: ammota@fc.ul.pt

Abstract

Background and Objective: Total Variation (TV) minimization algorithms have achieved great attention due to the virtue of decreasing noise while preserving edges. The purpose of this work is to implement and evaluate two TV minimization methods in 3D. Their performance is analyzed through 3D visualization of digital breast tomosynthesis (DBT) data with volume rendering.

Methods: Both filters were studied with real phantom and one clinical DBT data. One algorithm was applied sequentially to all slices and the other was applied to the entire volume at once. The suitable Lagrange multiplier used in each filter equation was studied to reach the minimum 3D TV and the maximum contrast-to-noise ratio (CNR). Imaging blur was measured at 0° and 90° using two disks with different diameters (0.5 mm and 5.0 mm) and equal thickness. The quality of unfiltered and filtered data was analyzed with volume rendering at 0° and 90° .

Results: For phantom data, with the sequential filter, a decrease of 25% in 3D TV value and an increase of 19% and 30% in CNR at 0° and 90° , respectively, were observed. When the filter is applied directly in 3D, TV value was reduced by 35% and an increase of 36% was achieved both for CNR at 0° and 90° . For the smaller disk, variations of 0% in width at half maximum (FWHM) at 0° and a decrease of about 2.5% for FWHM at 90° were observed for both filters. For the larger disk, there was a 2.5% increase in FWHM at 0° for both filters and a decrease of 6.28% and 1.69% in FWHM at 90° with the sequential filter and the 3D filter, respectively. When applied to clinical data, the performance of each filter was consistent with that obtained with the phantom.

Conclusions: Data analysis confirmed the relevance of these methods in improving quality of DBT images. Additionally, this type of 3D visualization showed that it may play an important complementary role in DBT imaging. It allows to visualize all DBT data at once and to analyze properly filters applied to all the three dimensions.

Concise Abstract. Total Variation (TV) minimization algorithms are one compressed sensing technique that has achieved great attention due to the virtue of decrease noise while preserve edges transitions. The purpose of this work is to solve the same TV minimization problem in DBT data, by studying two 3D filters. The obtained results were analyzed at 0° and 90° with a

3D visualization through volume rendering. The filters differ in their application. One considers a slice-by-slice optimization, sequentially traversing all slices of the data. The other considers the intensity values of adjacent slices to make this optimization on each voxel. The performance of each filter was also tested with a clinical case. The results obtained were very encouraging with a significantly increased contrast to noise ratio at 0° and 90° and a small reduction in blur at 90° (slight reduction of the out-of-plane artifact).

Keywords: Digital Breast Tomosynthesis; Total Variation minimization; volume rendering.

4.1. Introduction

Breast cancer alone accounts for 30% of all new cancer diagnoses in women and remains the second cause of death by cancer. Besides improvements in treatment, early detection has led to a reduction of approximately 40% in the death rate from this disease over the last three decades [1, 2]. This early detection is mostly done through screenings [3, 4]. Approved by the Food and Drug Administration less than ten years ago [5], Digital Breast Tomosynthesis (DBT) is firmly established and is an increasingly used imaging technique for breast cancer screening and diagnosis [6-10].

So far, DBT data is displayed one slice at a time or sequentially as a continuous cine loop [11]. This procedure hampers the judgment of potential lesions such as clusters of microcalcification, which can be spread across several slices and difficult to interpret in a two dimensional (2D) image [12]. In addition, because in one DBT exam there are about 30 times more images than in 2D Digital Mammography (DM), the mean reading time doubles for DBT when compared with 2D DM examination [13-15]. Fatigue caused by the analysis of large data sets can hamper the routine functions of a radiologist. Besides, the time required for each DBT data set evaluation is crucial both in clinical and screening environments, directly influencing the number of examinations interpreted [15]. Computer Aided Detection (CAD) systems developed for DBT have presented promising preliminary results regarding the reduction of this time [16, 17]. However, these systems are adapted to the 2D slice-by-slice visualization, which does not result immediately in a global inspection of the data. For example, in the case of clusters of microcalcifications, although decreasing the number of slices to be analyzed, it requires the radiologist to review multiple adjacent slices. In addition, these CAD systems, widely used in 2D DM, still present some controversy since they result in some false positives, sometimes leading to a more time consuming inspection [18].

Two other new approaches emerged in the visualization of DBT data: synthetic mammography build upon DBT data [19, 20] and thicker slabs obtained by combining several slices [21, 22]. The existence of synthetic mammography is very important because it allows a fair comparison with previous DM examinations and, in some situations, it could benefit from the CAD systems developed for 2D application. However, as DM, synthetic mammography still has the disadvantages of a 2D visualization. For example, in the case of a dense parenchyma pattern or microcalcification clusters, its clinical value is limited due to tissue overlapping (it is

used in combination with DBT slice-by-slice) [23]. On the other hand, the construction of slabs has demonstrated good results in terms of reducing time and false positives but originates a lower sensitivity [22].

A different type of visualization may play an important complementary role in breast cancer diagnosis [24]. Three dimensional (3D) volume rendering is the process of creating realistic computer-generated images of a 3D scene, yielding an improved depth perception [25]. 3D volume rendering proved to be useful in tomographic medical imaging modalities such as Computed Tomography (CT), Magnetic Resonance Imaging or Positron Emission Tomography [26-28]. There are some works mentioning a few aspects of 3D volume rendering for DBT [29-31] and its importance to detect clusters of microcalcifications [32]. As in CT, the 3D visualization integration in a 2D reading can be seen as a problem solver in some situations (for example, the necessity to better understand the shape of a cluster of microcalcifications) [33].

In DBT, a small number of low-dose projections acquired within a limited angular range are reconstructed to produce 3D data of breast tissue. Both the low dose per projection and the angular limit raise some complications in the reconstruction of DBT data. The first results in high noise level in the reconstructed images, while the second restricts the spatial resolution in the direction perpendicular to the detector plane (the z direction), resulting in out-of-plane blur. Visible in 2D slice-by-slice visualization, the out-of-plane artifact becomes even more evident in 3D volume rendering. Reducing noise without blurring or decreasing details and edge definition is a challenge in image processing. Total Variation (TV) minimization algorithms are efficient in reducing noise while preserving edges. For this reason, studies applying TV minimization to DBT data have grown significantly. Most works focus on the use of these algorithms in iterative reconstruction [34-41], with a few focusing on post-processing [42-44]. The Lagrange multiplier (λ) is an important parameter in TV minimization [42] since it controls image regularization, giving a trade-off between removing noise and preserving edges. For this reason, to find the appropriate λ value is extremely important in this type of regularization.

In this paper, two TV minimization algorithms based on [42, 44] were applied to DBT reconstructed data and their impact on the volume rendering visualization was analyzed. The first is a 2D TV minimization filter applied to a single slice [42], while the second is a 3D TV minimization filter applied to the entire volume and analyzed through a preliminary study in a 2D way (in-plane analysis) [44]. In this work, the 2D TV minimization filter was applied sequentially to all phantom slices and both the results, from 2D and 3D filters, were analyzed

based on volumetric rendering of DBT phantom data, considering their performance in z-direction. Despite the limited resolution in z, volume rendering is useful to visualize all DBT data at once (with particular emphasis on high intensity features, such as calcifications). In addition, as it results in a truly 3D visualization, it is an appropriate way to analyze filters applied to all three dimensions. Since the Lagrange multiplier is a decisive parameter in TV minimization, the suitable λ to be used was studied (in a first phase to reach the minimum 3D TV of the data and in a second stage to achieve the maximum contrast to noise value in a specific region). Quantitative and visual analyses were conducted between unfiltered and filtered rendered images and also between the two different TV minimization filters. The proposed algorithms were also tested with one clinical DBT data set.

4.2. Methods

4.2.1. Data acquisition and reconstruction

To mimic the breast tissue, an acrylic phantom made by us was scanned with a Siemens MAMMOMAT Inspiration system (Siemens AG, Healthcare Sector, Erlangen, Germany) installed in a clinical facility (Hospital da Luz S.A., Lisbon, Portugal). The phantom was acquired with 28 kVp and 47 mAs. To simulate high-density lesions, the phantom contains two columns of aluminum disks of different diameters and 1 mm thickness (Figure 4.1). For this study the first column was considered.



Figure 4.1. Acrylic phantom simulating breast tissue and lesions of high attenuation (aluminum disks of different diameters and 1 mm thickness). Diameter of the first column disks (top to bottom): 5 mm, 3 mm, 1 mm, 0.5 mm, 2 mm and 4 mm, respectively.

Additionally, one clinical DBT data set with a large calcification of an anonymous patient was selected from the clinical facility database. This clinical data set was acquired with 26 kVp and 34 mAs.

The acquired data were reconstructed with the manufacturer algorithm which uses Filtered Back Projection (FBP) with some post-processing to reduce artifacts and image blurring [45]. The reconstructions have voxel sizes of $0.085 \times 0.085 \times 1.0 \text{ mm}^3$. The algorithms under study were implemented using MATLAB R2016b and run on a computer containing 4 Intel® Xeon(R) CPU E5620 @ 2.40 GHz.

4.2.2. Data visualization

In volume rendering, changing the azimuth of a camera rotates its position around the focal point [46]. In this case, two different angle position were used to visualize the DBT data: 0° and 90° . The first corresponds to the visualization parallel to the detector plate (i.e. the projection is made on the xy planes along z) and the second is the visualization perpendicular to the detector plate (i.e. projection is made on the xz planes along y).

The visualization software was developed in C++ using the Visualization Toolkit library (VTK) version 7.1.0. [46, 47]. For a better visualization quality, voxels were made isotropic (with dimensions $0.085 \times 0.085 \times 0.085 \text{ mm}^3$) using the *Lanczos* function available in VTK (used by default) [46, 48].

The methodologies in study were analyzed using 3D volume rendering visualization with a ray casting algorithm and compositing technique. Different rendering parameters yield different images. For this reason, the parameters were fixed for all situations so that a correct comparison could be made.

4.2.3. Image analysis

For image analysis, quantitative and qualitative comparisons were performed between unfiltered and filtered rendered images at 0° and 90° . For phantom quantitative analysis, two figures of merit were used: Contrast to noise ratio (CNR) and full width at half maximum (FWHM) of a Gaussian curve fitted to the profile of two disks – 0.5 mm and 5.0 mm. For CNR, a region of interest (ROI) over the 5 mm disk and four other ROIs over the surrounding background were drawn. The CNR was calculated with *Eq. 4.1*, where μ_F and μ_B stand for the

mean pixel values in the ROI over the selected feature (5.0 mm disk) and the background, respectively; and σ_B stands for the mean of standard deviations in the background ROIs.

$$CNR = \frac{\mu_F - \mu_B}{\sigma_B} \quad Eq. 4.1$$

At 0°, the FWHM of the 0.5 mm disk was considered as an indicator of spatial resolution. At 90°, the FWHM of the 0.5 mm and 5 mm disks were considered as indicators of the blur in the z direction.

4.2.4. Formulation of TV minimization problem

The 3D TV values were calculated according to Eq. 4.2 where $u_{i,j,k}$ is the intensity value of voxel (i, j, k) , with $i = \{1, \dots, m\}$, $j = \{1, \dots, n\}$, $k = \{1, \dots, p\}$ and $m \times n \times p$ the data dimensions. Δ_x , Δ_y and Δ_z are discretizations of the horizontal (x), vertical (y) and perpendicular (z) derivatives, respectively.

$$TV_{3D}(u) = \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^p \sqrt{(\Delta_x u_{i,j,k})^2 + (\Delta_y u_{i,j,k})^2 + (\Delta_z u_{i,j,k})^2} \quad Eq. 4.2$$

The unconstrained TV minimization problem here addressed is based on ROF model [49] for Poisson noise [50, 51] and can be formulated as in Eq. 4.3. f and u are the original and denoised data, respectively, $TV(u)$ is the TV of denoised data and λ is the Lagrange multiplier, also called the regularization parameter.

$$\min_u \{TV(u) + \lambda(u - f \ln u)\} \quad Eq. 4.3$$

In Eq. 4.3, the first term corresponds to the regularization term (TV function) and the second one relates with the assumed noise model (fidelity term). In this way, λ controls data regularization, between removing noise and preserving information.

The problem represented in Eq. 4.3 can be solved through its Euler-Lagrange equation in finite data domain Ω ($\partial u / \partial n = 0$ on $\partial \Omega$, with n relative to noise). Eq. 4.4 stands for the 2D TV minimization filter and was obtained based on discretization of the corresponding Euler-Lagrange equation as in [42].

$$0 = \frac{\lambda}{u}(u - f) - \left[\Delta_x^- \frac{\Delta_x^+ u_{i,j}}{\sqrt{\varepsilon^2 + (\Delta_x^+ u_{i,j})^2 + (\Delta_y^0 u_{i,j})^2}} + \Delta_y^- \frac{\Delta_y^+ u_{i,j}}{\sqrt{\varepsilon^2 + (\Delta_x^0 u_{i,j})^2 + (\Delta_y^+ u_{i,j})^2}} \right] \quad Eq. 4.4$$

$u_{i,j}$ is the intensity value of voxel (i,j) at slice k . Each voxel to be denoised only depends on the intensity values of the surrounding voxels in the same slice (Figure 4.2 (a)). As this is a 3D study, Eq. 4.4 was applied sequentially to all slices ($k = \{1, \dots, p\}$).

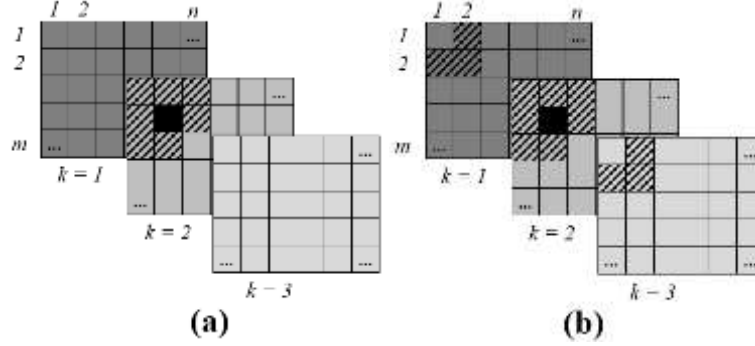


Figure 4.2. Illustrative scheme of the application of each filter in slice number two. (a) 2D TV minimization filter applied to all slices, one slice at a time. (b) 3D TV minimization filter.

In each slice, each voxel to be denoised with the 3D TV minimization filter is influenced by the intensity values of the surrounding voxels, including voxels in the neighboring slices (Figure 4.2 (b)). In this work, the discrete equation (Eq. 4.5) which corresponds to the Euler-Lagrange solution of the 3D TV minimization problem, presents some differences in relation to [44], specifically in the discretization of derivatives. These minor changes were achieved after some empirical studies where it was found that these combinations allow for a better balance between differences.

$$\begin{aligned}
 0 = \frac{\lambda}{u} (u - f) - & \left[\Delta_x^- \frac{\Delta_x^+ u_{i,j,k}}{\sqrt{\varepsilon^2 + (\Delta_x^+ u_{i,j,k})^2 + (\Delta_y^0 u_{i,j,k})^2 + (\Delta_z^0 u_{i,j,k})^2}} \right. \\
 & + \Delta_y^- \frac{\Delta_y^+ u_{i,j,k}}{\sqrt{\varepsilon^2 + (\Delta_x^0 u_{i,j,k})^2 + (\Delta_y^+ u_{i,j,k})^2 + (\Delta_z^0 u_{i,j,k})^2}} \\
 & \left. + \Delta_z^- \frac{\Delta_z^+ u_{i,j,k}}{\sqrt{\varepsilon^2 + (\Delta_x^0 u_{i,j,k})^2 + (\Delta_y^0 u_{i,j,k})^2 + (\Delta_z^+ u_{i,j,k})^2}} \right]
 \end{aligned} \tag{Eq. 4.5}$$

In Eq. 4.4 and Eq. 4.5, Δ_x^+ , Δ_x^- , Δ_y^+ , Δ_y^- , Δ_z^+ and Δ_z^- , denote forward (+) and backward (-) one-sided differences in x , y and z directions, respectively; Δ_x^0 , Δ_y^0 and Δ_z^0 indicate central difference in x , y and z directions, respectively; and $\varepsilon > 0$ is a small parameter introduced to remove the derivative singularity when u is locally constant.

These formulations are valid only for interior points, i.e., voxel (i,j,k) with $i = \{2, \dots, m-1\}$, $j = \{2, \dots, n-1\}$ and $k = \{2, \dots, p-1\}$, excluding borders. For the 2D problem (Eq. 4.4), boundary

conditions were defined as in [49]. For the 3D algorithm (Eq. 4.5), boundary conditions were defined as:

$$u_{1,j,k} = u_{2,j,k} \text{ and } u_{m,j,k} = u_{m-1,j,k} \text{ for } j = \{2, \dots, n-1\} \text{ and } k = \{2, \dots, p-1\}$$

$$u_{i,1,k} = u_{i,2,k} \text{ and } u_{i,n,k} = u_{i,n-1,k} \text{ for } i = \{2, \dots, m-1\} \text{ and } k = \{2, \dots, p-1\}$$

$$u_{i,j,1} = u_{i,j,2} \text{ and } u_{i,j,p} = u_{i,j,p-1} \text{ for } i = \{2, \dots, m-1\} \text{ and } j = \{2, \dots, n-1\}$$

$$u_{1,1,1} = u_{2,2,1}, u_{1,n,1} = u_{2,n-1,1}, u_{m,1,1} = u_{m-1,2,1}, u_{m,n,1} = u_{m-1,n-1,1},$$

$$u_{1,1,p} = u_{2,2,p}, u_{1,n,p} = u_{2,n-1,p}, u_{m,1,p} = u_{m-1,2,p}, u_{m,n,p} = u_{m-1,n-1,p}$$

The main differences between the previously studied algorithms [42, 44] and those applied in this study are summarized in the Table 4.1.

Table 4.1. Summary of the differences between the previous studies using the mentioned TV minimization algorithms [42, 44] and the studies developed in this work.

	Ref. [42]	TV min in all slices (Figure 4.2 (a))	Ref. [44]	TV min in the volume (Figure 4.2 (b))
Phantom	Mammographic Accreditation Phantom Model 156	Acrylic phantom simulating breast tissue and lesions of high attenuation	Mammographic Accreditation Phantom Model 156	Acrylic phantom simulating breast tissue and lesions of high attenuation
Reconstructed voxel (mm³)	0.34 x 0.34 x 1.0	0.085 x 0.085 x 1.0	0.34 x 0.34 x 1.0	0.085 x 0.085 x 1.0
Reconstruction algorithm	Algebraic Reconstruction Technique	Filtered Backprojection	Algebraic Reconstruction Technique	Filtered Backprojection
Implementation	<p>Ref. [42] was applied to a single slice of interest while, in this study, the algorithm of Figure 4.2 (a) was applied to all slices sequentially (is a 3D filter).</p> <p>A different combination of discrete derivatives was used in Eq. 4.5. In the square root, the combination in this study was: (+,0,0), (0,+,0) and (0,0,+) instead of (+,0,-), (-,+,0) and (0,-,+) used in the previous study [44]. With this combination there was an improvement of about 0.3% in 3D TV minimization.</p>			

4.2.5. Optimization of TV minimization filter

Filters' performance directly depend on the regularization parameter (λ) used in Eq. 4.4 and Eq. 4.5. λ allows to control the weight between the two terms in Eq. 4.3: Regularization Vs. Fidelity.

4.2.5.1. Comprehensive study of λ

There is one λ value for which the minimum 3D TV is obtained, while maintaining the fidelity of data. Although both the algorithms take into account the entire data (not only one slice), their

application is different (Figure 4.2). For this reason, minimum 3D TV values of filtered data will be different, and so the respective λ values ranging between [10, 250] were applied and the results obtained for both filters were analyzed.

The validity of 2D filter when applied to one single slice has already been carefully studied in [42] while preliminary results achieved with the 3D filter were presented in [44]. The aim of this comprehensive study was to validate both the algorithms under the modified conditions: the 2D algorithm considering all slices and the 3D filter after the introduced alterations. Additionally, it was also useful to find λ reference values for the next section.

4.2.5.2. λ vs. CNR_{max}(90°)

The λ value which allows the greatest minimization of 3D TV of the entire data may differ from the value which generates reduction of local blur or improvement of local CNR. In addition, as there is a minimization of TV and therefore noise and variability reduction, the impact of these filters in z-direction should be studied through image quality at 90°. For this reason, based on λ values responsible for the greatest minimization of 3D TV, a detailed study about CNR at 90° as a function of λ was conducted. Images generated by λ values allowing the highest CNR at 90° have been analyzed.

4.2.6. DBT clinical data

For the clinical quantitative analysis, CNR and FWHM were also obtained at 0° and 90°. CNR was calculated as in *Eq. 4.1*, with the feature of interest being the biggest calcification. The FWHM was obtained with a Gaussian curve fitted to the profile of the same calcification. It is known that the calcification diameter is approximately 3.6-3.8 mm at 0° (xy planes) and 8-11 mm in 90° (yz planes), respectively. As in phantom measures, FWHM value at 90° was considered as an indicator of the blur in z direction.

4.3. Results

Both filters were applied to the phantom with original dimensions of $1330 \times 1460 \times 42$ and voxel size $0.085 \times 0.085 \times 1.0$ mm³. 2D TV minimization algorithm took approximately 17 seconds to cover all slices and 3D TV minimization was done in approximately 28 seconds. The results obtained are shown in Sections 4.3.1-4.3.4.

4.3.1. Comprehensive study of λ vs. 3D TV values

The results for 3D TV values (calculated using *Eq. 4.2*) obtained with λ ranging between 10 and 250 (with a fixed step-size of 10) are presented in Figure 4.3 and Figure 4.4 for the 2D TV minimization in all slices and 3D TV minimization, respectively. For more detailed information, the λ value was varied with a step-size of 1 and the first filter was then applied with λ between 188 and 208 (Figure 4.3-right) and the second one with λ ranging from 215 and 235 (Figure 4.4-right).

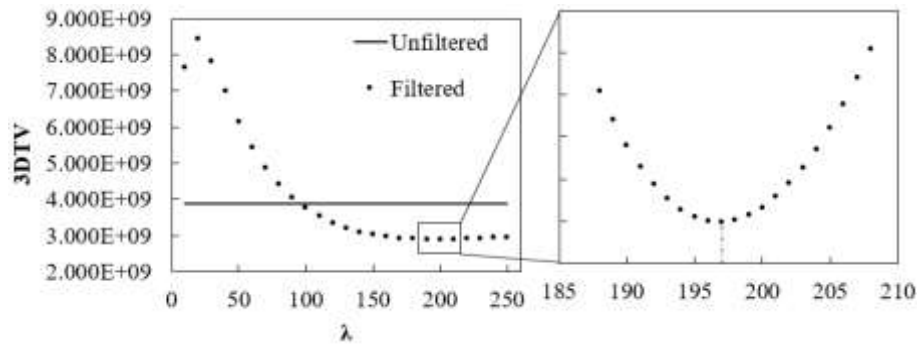


Figure 4.3. 3D TV values for unfiltered and filtered phantom data, with minimization of 2D TV in all slices, plotted as a function of λ . Right: zoom in on λ range where 3D TV is minimum.

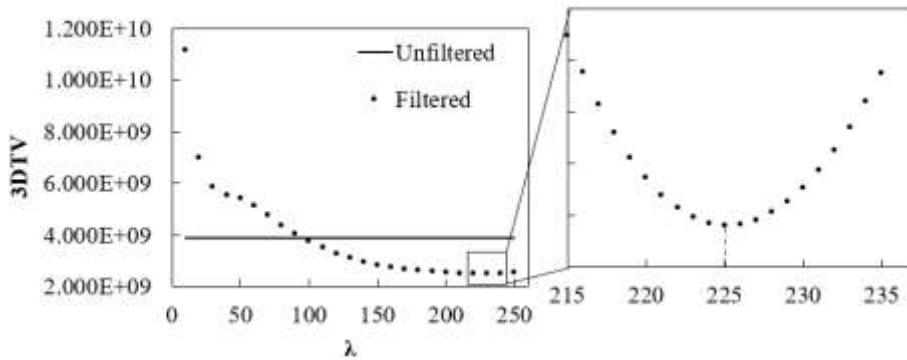


Figure 4.4. 3D TV values for unfiltered and filtered phantom data, with minimization of 3D TV, plotted as a function of λ . Right: zoom in on λ range where 3D TV is minimum.

The 3D TV value calculated for the unfiltered data and minimum 3D TV values obtained with each filter (and respective λ) are presented in Table 4.2.

Table 4.2. Summary of the results obtained for the 3D TV values of unfiltered and filtered phantom data with the two applications (minimization of 2D TV in all slices and minimization of 3D TV). It is also presented the variation in percentage between the unfiltered and filtered values.

	TV3D	λ	Δ (%)
Unfiltered	3.885E+09		
Filtered min2DTV all slices	2.892E+09	197	-25.57
Filtered min3DTV	2.519E+09	225	-35.17

4.3.2. λ vs. $\text{CNR}_{\text{max}}(90^\circ)$

From Table 4.2, $\lambda(\text{min2DTV-all slices})=197$ and $\lambda(\text{min3DTV})=225$ allow the lowest 3D TV value for each filter. For this reason, CNR calculated values in the 5 mm disk at 90° in the phantom filtered data as a function of λ values ranging between [155, 245] are presented in Figure 4.5.

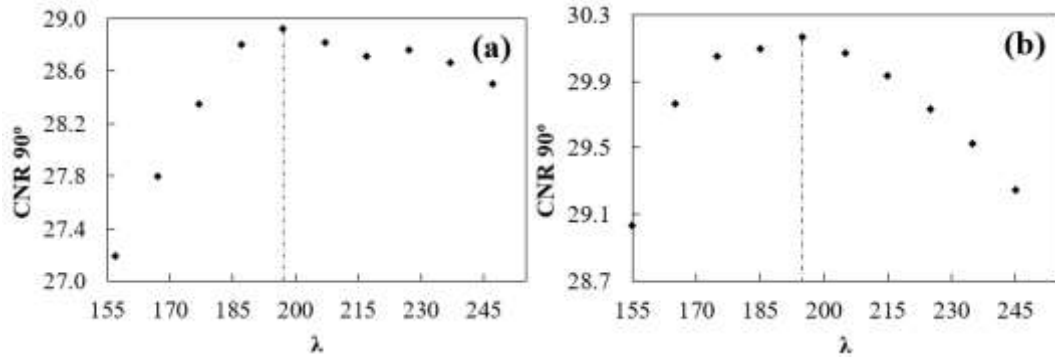


Figure 4.5. Values of CNR at 90° obtained with 2D TV minimization algorithm applied to all slices (a) and 3D TV minimization algorithm (b) as a function of λ .

4.3.3. Phantom analysis

From Figure 4.5 (a), for the 2D TV minimization algorithm, $\lambda=197$ was the value which allows the maximum CNR at 90° . On the other hand, $\lambda=195$ was the best choice for the 3D TV minimization algorithm (Figure 4.5 (b)). Table 4.3 summarizes the quantitative results obtained for the filtered data at 90° using these λ values.

Table 4.3. Results obtained for CNR, FWHM_{0.5mm} and FWHM_{5.0mm} values of unfiltered and filtered phantom data at 90° . Each filtered data was obtained using λ corresponding to the maximum CNR at 90° ($\lambda=197$ for the 2D TV minimization in all slices and $\lambda=195$ for the 3D TV minimization). It is also presented the variation in percentage between the unfiltered and filtered values.

	90°					
	CNR_{max}		$\text{FWHM}_{0.5\text{mm}}$ (mm)		$\text{FWHM}_{5.0\text{mm}}$ (mm)	
Unfiltered	22.106	Δ (%)	2.594	Δ (%)	3.100	Δ (%)
Filtered min2DTV all slices	28.924	30.84	2.517	-2.97	2.906	-6.28
Filtered min3DTV	30.167	36.47	2.533	-2.36	3.048	-1.69

As DBT data and algorithms are 3D, besides analysis in z-direction, it is important to also ensure quality at 0° . In this way, the quantitative values obtained in the xy reconstruction plane are presented in Table 4.4.

Table 4.4. Results obtained for CNR, FWHM_{0.5mm} and FWHM_{5.0mm} values of unfiltered and filtered phantom data at 0° . Each filtered data was obtained using λ corresponding to the maximum CNR at 90° ($\lambda=197$ for the 2D TV minimization in all

slices and $\lambda=195$ for the 3D TV minimization). It is also presented the variation in percentage between the unfiltered and filtered values.

0°						
	CNR_{\max}		$\text{FWHM}_{0.5\text{mm}} (\text{mm})$		$\text{FWHM}_{5.0\text{mm}} (\text{mm})$	
Unfiltered	17.427	$\Delta (\%)$	0.4957	$\Delta (\%)$	4.479	$\Delta (\%)$
Filtered min2DTV all slices	20.822	19.48	0.4956	-0.02	4.595	2.58
Filtered min3DTV	23.761	36.35	0.4962	0.09	4.598	2.65

Images of results achieved with volume rendering at 0° of the 0.5 mm and 5 mm disks are presented in Figure 4.6 and images at 90° are shown in Figure 4.7. In both Figure 4.6 and Figure 4.7, the first column (a1 and a2) represent unfiltered disks, the second column (b1 and b2) presents filtered disks with 2D TV minimization in all slices, and the last column (c1 and c2) shows the results obtained with 3D TV filter.

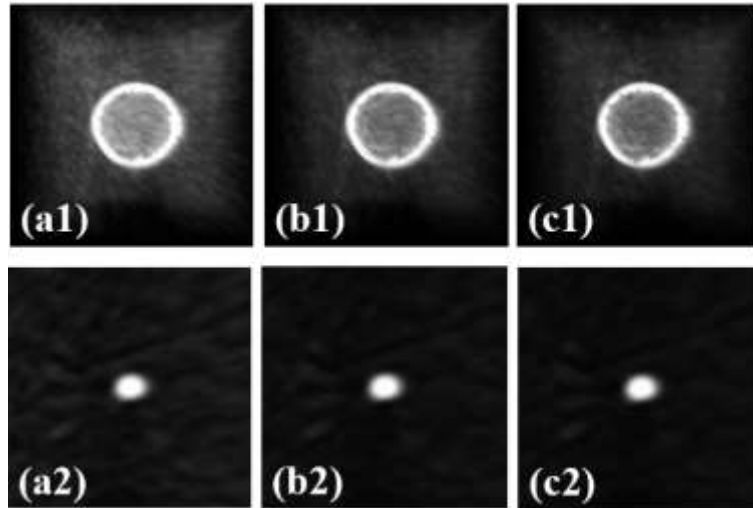


Figure 4.6. Volume rendering images at 0° for 5 mm disk (top row) and 0.5 mm disk (bottom row) obtained for the unfiltered (a1 and a2) and filtered data with 2D TV minimization filter applied to all slices – $\lambda=197$ (b1 and b2) and 3D TV minimization filter – $\lambda=195$ (c1 and c2).

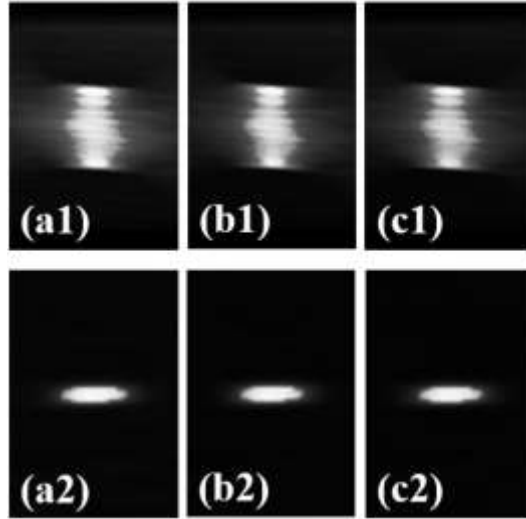


Figure 4.7. Volume rendering images at 90° for 5 mm disk (top row) and 0.5 mm disk (bottom row) obtained for the unfiltered (a1 and a2) and filtered data with 2D TV minimization filter applied to all slices – $\lambda=197$ (b1 and b2) and 3D TV minimization filter – $\lambda=195$ (c1 and c2).

4.3.4. Clinical data analysis

The results for 3D TV values obtained for clinical data with λ ranging between 10 and 200 are presented in Figure 4.8 and Figure 4.9 for the 2D TV minimization in all slices and 3D TV minimization, respectively. For further analysis, 3D TV values of filtered clinical data were obtained with λ ranging from 124 and 144 for the first filter (Figure 4.8-right) and λ between 135 and 155 for the second (Figure 4.9-right).

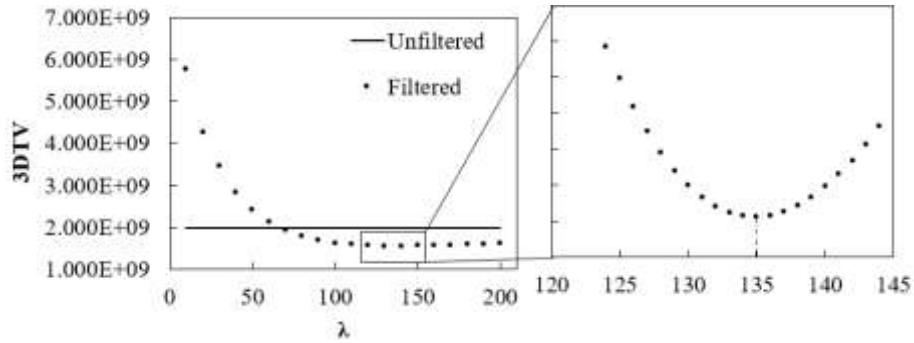


Figure 4.8. 3D TV values for unfiltered and filtered clinical data, with minimization of 2D TV in all slices, plotted as a function of λ . Right: zoom in on λ range where 3D TV is minimum.

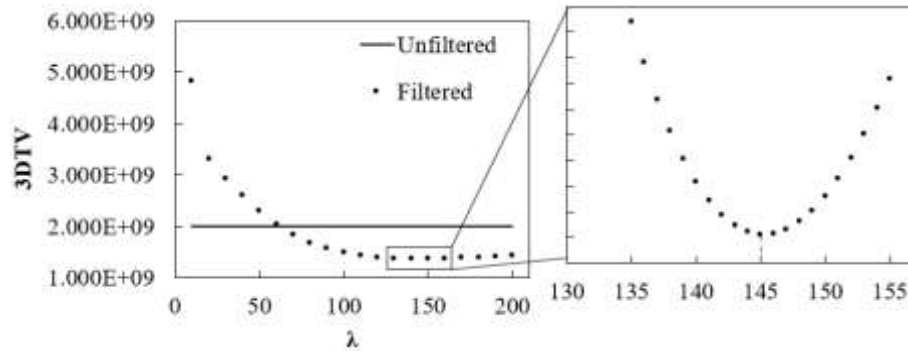


Figure 4.9. 3D TV values for unfiltered and filtered clinical data, with minimization of 3D TV, plotted as a function of λ . Right: zoom in on λ range where 3D TV is minimum.

The 3D TV value calculated for the unfiltered clinical data and the minimum 3D TV values obtained after the application of each filter (and respective λ) are summarized in Table 4.5.

Table 4.5. Summary of the results obtained for the 3D TV values of unfiltered and filtered clinical data with the two applications (minimization of 2D TV in all slices and minimization of 3D TV). It is also presented the variation in percentage between the unfiltered and filtered values.

	TV3D	λ	Δ (%)
Unfiltered	1.995E+09		
Filtered min2DTV all slices	1.556E+09	135	-22.02
Filtered min3DTV	1.371E+09	145	-31.29

In Table 4.6 are presented the CNR and FWHM values measured, at 0° and 90° , before and after the application of each filter using the respective λ value that allowed to obtain the minimum 3D TV value. In Figure 4.10, volume rendering of the results obtained with clinical data at 0° and 90° are illustrated.

Table 4.6. Results obtained for CNR and FWHM_{Calc} values of unfiltered and filtered clinical data at 0° and 90° . Each filtered data was obtained using λ corresponding to the minimum 3D TV ($\lambda=135$ for the 2D TV minimization in all slices and $\lambda=145$ for the 3D TV minimization). It is also presented the variation in percentage between the unfiltered and filtered values.

0°				
	CNR		FWHM _{Calc} (mm)	
Unfiltered	20.483	Δ (%)	3.660	Δ (%)
Filtered min2DTV all slices	28.490	39.09	3.657	-0.06
Filtered min3DTV	30.607	49.43	3.667	0.20
90°				
	CNR		FWHM _{Calc} (mm)	
Unfiltered	18.900	Δ (%)	9.865	Δ (%)
Filtered min2DTV all slices	22.756	20.40	9.318	-5.55
Filtered min3DTV	23.701	25.40	9.501	-3.69

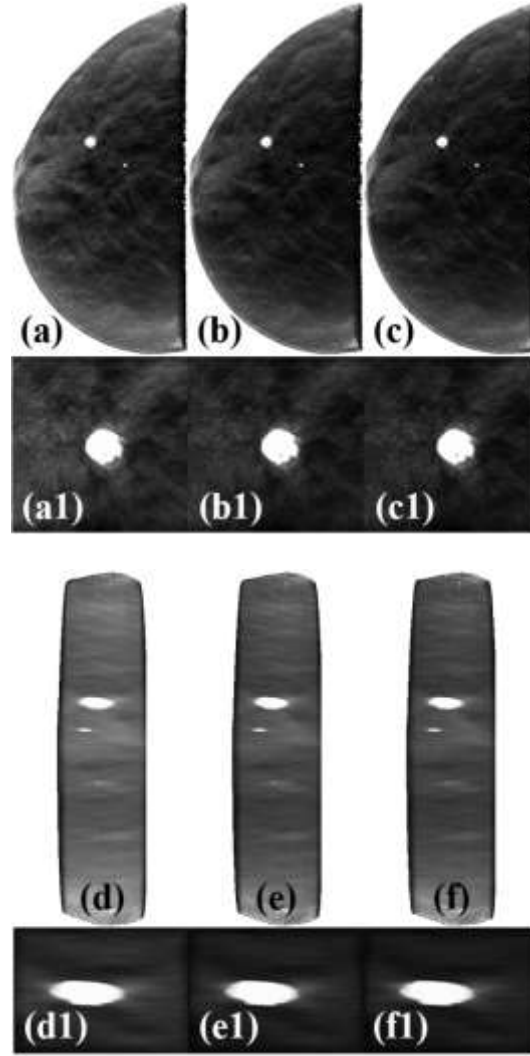


Figure 4.10. Volume rendering at 0° (a, b and c) and 90° (d, e and f) of clinical data. (a1, b1, c1, d1, e1 and f1) represent the zoom in of the biggest calcification. Unfiltered data: a (a1) and d (d1); filtered data with 2D TV minimization: b (b1) and e (e1); and data obtained with 3D TV minimization filter: c (c1) and f (f1).

4.4. Discussion

The application of this type of filters after reconstruction is a straightforward approach and it is not time-consuming, unlike iterative methods. In addition, these filters are applied to all dimensions of DBT data and the visualization of obtained results through 3D volume rendering allows a better analysis of the effects in all directions. Here, two TV minimization filters were optimized considering the three dimensions and our study focused on: the λ value that solves the 3D TV minimization problem, validating the algorithms and the λ value which gives the maximum CNR in a high intensity region at 90° (without compromising or even improving the z-blur).

Considering the results obtained with the phantom, we observe that both filters had more influence in CNR values than on FWHM, at 0° and 90° . Regardless of the filter, the main objectives of minimizing 3D TV and increasing the CNR, without affecting the blur in the z -direction, were achieved. Taking into account Figure 4.3 and Figure 4.4, filtered 3D TV values are smaller than the 3D TV value of unfiltered data for λ values greater than 100. λ works as a weighting factor in *Eq. 4.3*, which means, it should ensure that the fidelity term has enough weight to keep fundamental information in the filtered data. For 3D TV values achieved with λ inferior to 100, the fidelity term has not sufficient influence on the results as there is too much regularization and unwanted artifacts are introduced. Therefore, in both cases, only results obtained with λ values greater than 100 were considered as valid. It is observed that both the algorithms presented the expected behavior for this problem, achieving a minimum value of 3D TV, with a significant reduction when compared with unfiltered data. The 2D minimization filter applied to all slices allowed a 25.57% decrease with $\lambda=197$, while the 3D filter achieved a 35.17% reduction in the 3D TV value with $\lambda=225$ (Table 4.2). As 3D TV considers alterations in x , y and z , the algorithm which is truly applied in (taking into account adjacent slices - Figure 4.2 (b)), results in the greatest reduction.

The λ value which allowed the maximization of CNR at 90° was the same for the sequentially applied 2D minimization filter (197), while for the 3D minimization it was changed to 195 (Figure 4.5). With these λ values, a significant improvement in CNR at 90° was achieved: 30.84% and 36.47% for the 2D and 3D filters, respectively (Table 4.3). These quantitative values resulted in smoother rendered images in Figure 4.7 (b1) and (c1) when compared with Figure 4.7 (a1). Blur in z -direction, measured through FWHM of 0.5 mm and 5 mm disks at 90° , has not increased in any case, achieving a reduction of 6.28% for 5 mm disk when filtered with the 2D algorithm slice-by-slice (Table 4.3). This effect is visible through a slight decrease in disk thickness in Figure 4.7 (b1). For the 0.5 mm disk (Figure 4.7 (a2) to (c2)), the differences are not visually evident (with variations of about 2% to 3% in FWHM). 3D TV minimization filter presented better results for the increase in CNR (and resulting noise reduction) but 2D minimization filter resulted in a greater reduction in z -blur. This is due to its slice-by-slice application, where it does not spread information between adjacent slices to make data smoother. Although the actual thickness of the disks is 1 mm, both presented higher values (Table 4.3) and the larger the disk, the greater the contamination in z . Despite reducing this contamination, this type of filter does not solve this problem. Deconvolution methods may be

a good option but are quite sensitive to noise. As these TV minimization filters showed a significant increase in CNR (reducing noise), a combination of both approaches can yield interesting results that should be studied in the future.

As pointed out in this paper, all data dimensions are important. In this way, it is essential to also ascertain image quality at 0° , i.e., in xy (usual 2D visualization plane). From Table 4.4, we observe the results obtained at 0° are consistent with those at 90° . Both filters avoid FWHM deterioration with a variation of about 0% for the smallest disk and 2.5% for the largest disk. The 2D minimization filter allowed a 19.48% increase in CNR, while 3D filter again showed a better result with an increase of 36.25% (similar to 90°). This improvement in CNR is quite visible in Figure 4.6 (a1) to (c1) through smoother images with a crescent degree from left (Figure 4.6 (a1)) to right (Figure 4.6 (c1)). The small variations in FWHM resulted in preservation of the edges and alterations not visible to naked eye in the results of Figure 4.6, both for the larger disk (Figure 4.6 (a1) to (c1)) and for the smallest (Figure 4.6 (a2) to (c2)).

Considering clinical data, results obtained with λ values above 70 were considered valid for both algorithms (Figure 4.8 and Figure 4.9). This lower threshold (when compared to phantom) varies according to σ_B of the unfiltered data. As this value is higher for clinical data (which contain much more anatomical information), the minimum value of λ must be lower [42]. With $\lambda = 135$, 2D filter achieved a 22.02% minimization on 3D TV. The 3D filter reached a 3D TV minimum (decrease of 31.29%) with $\lambda=145$ (Table 4.5). Besides, according to Table 4.6, the results of the metrics studied at 0° and 90° are in agreement with those obtained for the phantom. At 0° , the CNR has here a more significant increase: 39.09% and 49.43% for the 2D and 3D filters, respectively. At 90° , the variations are of a similar order of magnitude: 20.40% and 25.40% for each case, respectively. CNR measures in clinical data should be interpreted with caution because clinical background often considers anatomical noise, which may influence the results. On the other hand, taking into account the calcification in this case, also the variations obtained for FWHM are similar to those of the phantom. At 0° , these variations are numerically small ($\sim 0\%$) for both filters. At 90° , the 2D filter achieves once again a greater blur reduction (5.55%) than 3D filter (3.69%). Images from Figure 4.10 are in accordance with these quantitative results. Areas of higher noise (clearly visible in the zoom of Figure 4.10 (a1) and (d1)) become much softer with the application of any of the filters (Figure 4.10 (b1, c1, e1, f1)). Calcification edges were preserved at 0° and its spread was reduced at 90° (Figure 4.10 (e1)).

As mentioned before, different rendering parameters such as voxel size in z, sample distance or transfer functions can influence the obtained results. Therefore, the presented results are compared only with each other or with results obtained under the same conditions. In addition to visualization parameters, also acquisition parameters such as mAs (which produces data with different noise levels) and kVp (which produces data with different contrast levels) can affect the results. Despite these limitations, the presented and validated methodologies are reproducible in other scenarios.

We think it is important to compare these results obtained after reconstruction with FBP with DBT data reconstructed with iterative algorithms, which have shown interesting results in terms of reducing the out-of-plane artifact.

4.5. Conclusions

To solve the same TV minimization problem in DBT data, two 3D filters were studied. The obtained results were analyzed at two angles (0° and 90°) with a 3D visualization through volume rendering. The filters differ in their application. One considers a slice-by-slice optimization, sequentially traversing all slices of the data. The other considers the intensity values of adjacent slices to make this optimization on each voxel.

For the rendered phantom images, the suitable Lagrange multiplier (λ) to be used in the discrete equation of each filter has been studied with two objectives: first, 3D TV minimization (validating the algorithms) and second, CNR maximization at 90°, without increasing the blur in z. Analysis at other visualization angles (such as at 45°) should be considered in future work.

The performance of each filter was also tested with a clinical case containing a calcification. The results obtained for the phantom and the clinical data are very encouraging with a significantly increased CNR at 0 ° and 90 ° and a small reduction in blur at 90 ° (slight reduction of the out-of-plane artifact). In addition, both approaches were applied in a considerably short time, taking into account the number of voxels in question.

This study presents the particularity of two 3D filters and analyzes the performance via 3D visualization. In this way, it is possible to get a sense of the impact of these algorithms on the data in a direct way, by visualizing the DBT data at once from several angles.

ACKNOWLEDGMENTS

This work was supported by Universidade de Lisboa (PhD grant) and Fundação para a Ciência e Tecnologia – Portugal (Grant No. SFRH/BD/135733/2018 and FCT-IBEB Strategic Project UID/BIO/00645/2013).

References

1. Ferlay, J., et al., *Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018*. European Journal of Cancer, 2018. **103**: p. 356-387.
2. Siegel, R.L., K.D. Miller, and A. Jemal, *Cancer statistics, 2019*. CA Cancer J Clin, 2019. **69**(1): p. 7-34.
3. Berry, D.A., et al., *Effect of screening and adjuvant therapy on mortality from breast cancer*. N Engl J Med, 2005. **353**(17): p. 1784-92.
4. Independent UK Panel on Breast Cancer Screening, *The benefits and harms of breast cancer screening: an independent review*. The Lancet, 2012. **380**(9855): p. 1778-1786.
5. Food and Drug Administration (FDA) U.S. . *Accreditation/Certification for Facilities Utilizing a DBT System with FFDM Images*. 2011 [cited 2020 February]; Available from: <http://www.fda.gov/Radiation-EmittingProducts/MammographyQualityStandardsActandProgram/FacilityCertificationandInspection/ucm114148.htm>.
6. Gennaro, G., et al., *Digital breast tomosynthesis versus digital mammography: a clinical performance study*. Eur Radiol, 2010. **20**(7): p. 1545-1553.
7. Brandt, K.R., et al., *Can Digital Breast Tomosynthesis Replace Conventional Diagnostic Mammography Views for Screening Recalls Without Calcifications? A Comparison Study in a Simulated Clinical Setting*. American Journal of Roentgenology, 2013. **200**(2): p. 291-298.
8. Bonafede, M.M., et al., *Value analysis of digital breast tomosynthesis for breast cancer screening in a commercially-insured US population*. ClinicoEconomics and Outcomes Research: CEOR, 2015. **7**: p. 53-63.
9. Gao, Y., et al., *Digital Breast Tomosynthesis Practice Patterns Following 2011 FDA Approval: A Survey of Breast Imaging Radiologists*. Acad Radiol, 2017. **24**(8): p. 947-953.
10. Destounis, S., A. Santacroce, and A. Arieno, *DBT as a Screening Tool and a Diagnostic Tool*. Current Breast Cancer Reports, 2017. **9**(4): p. 264-271.
11. Sechopoulos, I., *A review of breast tomosynthesis. Part I. The image acquisition process*. Med Phys, 2013. **40**(1): p. 014301.
12. Samala, R.K., et al., *Digital breast tomosynthesis: Computer-aided detection of clustered microcalcifications on planar projection images*. Phys Med Biol, 2014. **59**(23): p. 7457-7477.
13. Good, W.F., et al., *Digital breast tomosynthesis: a pilot observer study*. AJR Am J Roentgenol, 2008. **190**(4): p. 865-9.
14. Gur, D., et al., *Digital breast tomosynthesis: observer performance study*. AJR Am J Roentgenol, 2009. **193**(2): p. 586-91.
15. Caumo, F., et al., *Digital Breast Tomosynthesis with Synthesized Two-Dimensional Images versus Full-Field Digital Mammography for Population Screening: Outcomes from the Verona Screening Program*. Radiology, 2018. **287**(1): p. 37-46.
16. Benedikt, R.A., et al., *Concurrent Computer-Aided Detection Improves Reading Time of Digital Breast Tomosynthesis and Maintains Interpretation Performance in a Multireader Multicase Study*. American Journal of Roentgenology, 2017. **210**(3): p. 685-694.
17. Balleyguier, C., et al., *Improving digital breast tomosynthesis reading time: A pilot multi-reader, multi-case study using concurrent Computer-Aided Detection (CAD)*. Eur J Radiol, 2017. **97**: p. 83-89.

18. Tchou, P.M., et al., *Interpretation Time of Computer-aided Detection at Screening Mammography*. Radiology, 2010. **257**(1): p. 40-46.
19. Hofvind, S., et al., *Digital Breast Tomosynthesis and Synthetic 2D Mammography versus Digital Mammography: Evaluation in a Population-based Screening Program*. Radiology, 2018. **287**(3): p. 787-794.
20. Simon, K., et al., *Accuracy of Synthetic 2D Mammography Compared With Conventional 2D Digital Mammography Obtained With 3D Tomosynthesis*. American Journal of Roentgenology, 2019. **212**(6): p. 1406-1411.
21. van Schie, G., et al., *Mass detection in reconstructed digital breast tomosynthesis volumes with a computer-aided detection system trained on 2D mammograms*. Med Phys, 2013. **40**(4): p. 041902.
22. Iotti, V., et al., *Comparing two visualization protocols for tomosynthesis in screening: specificity and sensitivity of slabs versus planes plus slabs*. Eur Radiol, 2019. **29**(7): p. 3802-3811.
23. Petropoulos, A.E., et al., *Quantitative assessment of microcalcification cluster image quality in digital breast tomosynthesis, 2-dimensional and synthetic mammography*. Medical & Biological Engineering & Computing, 2020. **58**(1): p. 187-209.
24. Venson, J.E., et al., *A Case-Based Study with Radiologists Performing Diagnosis Tasks in Virtual Reality*. Stud Health Technol Inform., 2017. **245**: p. 244-248.
25. Suetens, P., *Medical image analysis*, in *Fundamentals of Medical Imaging* 2009, Cambridge University Press: New York. p. 159-189.
26. O'Connell, A., et al., *Cone-Beam CT for Breast Imaging: Radiation Dose, Breast Coverage, and Image Quality*. American Journal of Roentgenology, 2010. **195**(2): p. 496-509.
27. Song, H., X. Cui, and F. Sun, *Breast Tissue 3D Segmentation and Visualization on MRI*. International Journal of Biomedical Imaging, 2013. **2013**: p. 8.
28. Jung, Y., et al., *Occlusion and Slice-Based Volume Rendering Augmentation for PET-CT*. IEEE Journal of Biomedical and Health Informatics, 2017. **21**(4): p. 1005-1014.
29. Alyassin, A.M. *Automatic transfer function generation for volume rendering of high-resolution x-ray 3D digital mammography images*. in *Medical Imaging 2002*. 2002. SPIE.
30. Alyassin, A.M., et al., *3D Visualization of X-ray Tomosynthesis Digital Mammography Data: Preference Study*, in *Digital Mammography: IWDM 2002 — 6th International Workshop on Digital Mammography*, H.-O. Peitgen, Editor 2003, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 507-509.
31. Dharanija, R. and T. Rajalakshmi. *A Conjunct Analysis for Breast Cancer Detection by Volume Rendering of Low Dosage Three Dimensional Mammogram*. in *Progress In Electromagnetics Research Symposium Proceedings*. 2011. China.
32. Jerebko, A., et al. *3D rendering methods for visualization of clusters of calcifications in digital breast tomosynthesis: a feasibility study*. in *ECR 2011*. 2011. Vienna, Austria.
33. Neri, E., et al., *The second ESGAR consensus statement on CT colonography*. Eur Radiol, 2013. **23**(3): p. 720-729.
34. Sidky, E.Y., et al., *Enhanced imaging of microcalcifications in digital breast tomosynthesis through improved image-reconstruction algorithms*. Med Phys, 2009. **36**(11): p. 4920-32.
35. Ertas, M., et al., *Digital breast tomosynthesis image reconstruction using 2D and 3D total variation minimization*. Biomedical engineering online, 2013. **12**: p. 112-112.
36. Seyyedi, S., et al., *An object-oriented simulator for 3D digital breast tomosynthesis imaging system*. Comput Math Methods Med, 2013. **2013**: p. 250689-250689.
37. Michielsen, K., et al., *Design of a model observer to evaluate calcification detectability in breast tomosynthesis and application to smoothing prior optimization*. Med Phys, 2016. **43**(12): p. 6577-6587.
38. Liu, Y., et al., *An adaptive multiscale anisotropic diffusion regularized image reconstruction method for digital breast tomosynthesis*. Australasian Physical & Engineering Sciences in Medicine, 2018. **41**(4): p. 993-1008.
39. Gomi, T. and Y. Koibuchi, *Use of a Total Variation Minimization Iterative Reconstruction Algorithm to Evaluate Reduced Projections during Digital Breast Tomosynthesis*. BioMed Research International, 2018. **2018**: p. 5239082-5239082.

40. Garrett, J.W., et al., *Reduced anatomical clutter in digital breast tomosynthesis with statistical iterative reconstruction*. Med Phys, 2018. **45**(5): p. 2009-2022.
41. Krammer, J., et al., *Evaluation of a new image reconstruction method for digital breast tomosynthesis: effects on the visibility of breast lesions and breast density*. The British Journal of Radiology, 2019. **92**(1103).
42. Mota, A.M., et al., *Total variation minimization filter for DBT imaging*. Med Phys, 2015. **42**(6Part1): p. 2827-2836.
43. Mota, A.M., et al. *An iterative algorithm for Total Variation minimization in DBT imaging*. in *VipIMAGE 2015*, . 2015. Tenerife, Spain: CRC Press.
44. Mota, A.M., et al. *3D Total Variation Minimization Filter for Breast Tomosynthesis Imaging*. 2016. Malmö, Sweden: Springer International Publishing.
45. Siemens. *MAMMOMAT Inspiration - Tomosynthesis Option*. 2015 [cited 2020 February]; Available from: https://www.accessdata.fda.gov/cdrh_docs/pdf14/P140011c.pdf.
46. Schroeder, W., K. Martin, and B. Lorensen, *The Visualization Toolkit: An Object-oriented Approach to 3D Graphics*. 4rd ed2006, USA: Kitware.
47. VTK. *Visualization Toolkit - VTK*. 2020 [cited 2020 February]; Available from: <http://www.vtk.org/>.
48. VTK - Interpolators. *Visualization Toolkit - VTK - Interpolators*. [cited 2020 February]; Available from: https://vtk.org/Wiki/VTK/Image_Interpolators.
49. Rudin, L.I., S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*. Physica D: Nonlinear Phenomena, 1992. **60**(1): p. 259-268.
50. Le, T., R. Chartrand, and T.J. Asaki, *A Variational Approach to Reconstructing Images Corrupted by Poisson Noise*. Journal of Mathematical Imaging and Vision, 2007. **27**(3): p. 257-263.
51. Sawatzky, A., et al. *Total Variation Processing of Images with Poisson Statistics*. 2009. Berlin, Heidelberg: Springer Berlin Heidelberg.

An enhanced visualization of DBT imaging using blind deconvolution and total variation minimization regularization

5

Ana M. Mota^{a*}, Matthew J. Clarkson^b, Pedro Almeida^a and Nuno Matela^a, "An Enhanced Visualization of DBT Imaging Using Blind Deconvolution and Total Variation Minimization Regularization," in *IEEE Transactions on Medical Imaging*, vol. 39, no. 12, pp. 4094-4101, Dec. 2020, doi: 10.1109/TMI.2020.3013107.

IEEE Transactions on Medical Imaging is a peer-review journal which have an impact factor of **10.048** and is ranked in the **Q1** in the domains: "Computer science – interdisciplinary applications", "Biomedical Engineering", "Electrical & Electronic Engineering", "Imaging science & photographic technology" and "Radiology, Nuclear medicine & Medical imaging".

^aFaculdade de Ciências, Instituto de Biofísica e Engenharia Biomédica, Universidade de Lisboa, Campo Grande 1749-016 Lisboa, Portugal;

^bDepartment of Medical Physics and Biomedical Engineering and the Centre for Medical Image Computing, University College London, London, UK.

* Corresponding author: ammota@fc.ul.pt

Abstract

Digital Breast Tomosynthesis (DBT) presents out-of-plane artifacts caused by features of high intensity. Given observed data and knowledge about the point spread function (PSF), deconvolution techniques recover data from a blurred version. However, a correct PSF is difficult to achieve and these methods amplify noise. When no information is available about the PSF, blind deconvolution can be used. Additionally, Total Variation (TV) minimization algorithms have achieved great success due to its virtue of preserving edges while reducing image noise. This work presents a novel approach in DBT through the study of out-of-plane artifacts using blind deconvolution and noise regularization based on TV minimization. Gradient information was also included. The methodology was tested using real phantom data and one clinical data set. The results were investigated using conventional 2D slice-by-slice visualization and 3D volume rendering. For the 2D analysis, the artifact spread function (ASF) and Full Width at Half Maximum ($\text{FWHMM}_{\text{ASF}}$) of the ASF were considered. The 3D quantitative analysis was based on the FWHM of disks profiles at 90° , noise and signal to noise ratio (SNR) at 0° and 90° . A marked visual decrease of the artifact with reductions of FWHM_{ASF} (2D) and FWHM_{90° (volume rendering) of 23.8% and 23.6%, respectively, was observed. Although there was an expected increase in noise level, SNR values were preserved after deconvolution. Regardless of the methodology and visualization approach, the objective of reducing the out-of-plane artifact was accomplished. Both for the phantom and clinical case, the artifact reduction in the z was markedly visible.

Index Terms: Blind deconvolution, breast tomosynthesis, data visualization, inverse problems, total variation minimization.

5.1. Introduction

Digital Breast Tomosynthesis (DBT) has consolidated its position as an imaging technique used for breast cancer screening and diagnosis [1-5]. In DBT, a small number of low-dose projections acquired within a limited angular range are reconstructed to produce a three dimensional (3D) image of breast tissue. In the reconstruction, this angular limit restricts the spatial resolution in the direction perpendicular to the detector plane (z-direction), resulting in out-of-plane artifacts (blur in the z-direction). These artifacts are mostly produced by structures of high density that result in features of high attenuation and intensity (brighter) in DBT images. Examples are calcifications, biopsy needles and localization wires. The production of these artifacts could potentially obscure breast lesions and would limit the usability of DBT in interventional procedures. This is particularly important since DBT has a crucial role in guided biopsies for lesions such as architectural distortions (which can be occult in mammography and ultrasound images) [6, 7].

Currently, DBT images are analyzed with a two dimensional (2D) slice-by-slice visualization [8]. DBT presents, on average, sixty 2D slices per exam, which results in time-consuming analysis both in screening and daily clinical use [9-11]. Some studies have shown promising results in an attempt to reduce this time of analysis by highlighting some slices considered of interest and thus decreasing the number of slices to be analyzed [12, 13]. However, these approaches result in an increase in false positives, requiring the radiologist to review multiple adjacent slices, sometimes leading to a more time consuming inspection [14]. Volume rendering is a 3D visualization approach which can display data from any angle, resulting in an immediate global inspection [15]. Therefore, DBT visualization by volume rendering may represent a complementary option in the analysis of DBT examinations [16] because it provides an understanding of the underlying data at once. In 2D visualization through xy planes along the z-direction, the out-of-plane artifact makes the bright structures visible in the in-focus plane and replicated with lesser intensity to the underlying planes. In volume rendering this type of artifact remains quite evident with a noticeable blur at 90° (visualization perpendicular to detector plate).

As the out-of-plane artifact is a current drawback in DBT imaging, there are several works addressing this topic. Most of them are focused on addressing this problem by filtering the projections [17-20], i.e. during reconstruction, with a few focusing on post-processing [21-23]. However, even with some of these techniques already implemented by the manufacturers [24],

the artifact remains quite visible, with the characterization of calcifications and their morphology greatly affected by it [25].

The general imaging process can be denoted by the expression:

$$f = h \otimes u + n \quad \text{Eq. 5.1}$$

where f and u are the output (blurred) and the ideal data, respectively, h represents the degradation function and n is the noise model. In the spatial domain, h is referred to as the point spread function (PSF), which is responsible for blurring u . Deconvolution is an inverse filtering process in which the effects of convolution by a PSF (blurring) should be inverted [26]. Given f and some knowledge about h (PSF) and n , the goal of deconvolution methods is to obtain an estimate of the original data, u [27] using fast-Fourier transform algorithms [28].

One of the most difficult problems in image restoration is obtaining an adequate estimate of the PSF to use in deconvolution algorithms. The availability of suitable imaging equipment and phantoms which allow the correct determination of the PSF is limited [27]. When no information is available about the PSF, blind deconvolution can be used. The goal of blind deconvolution is to recover u and the PSF from f [28]. On the other hand, recovering the original data by inverse filtering involves noise amplification and DBT data are already noisy by nature due to the acquisition of low dose projections. Reducing noise without blurring or decreasing details and edge definition is a challenge in image processing. Total variation (TV) based minimization algorithms have achieved great success due to their ability to preserving edges while reducing image noise. TV is a quantity that characterizes how smoothly the intensity of an image is changing and it increases significantly in the presence of noise. Studies applying TV minimization to DBT data have grown significantly [29-35].

In this paper, a methodology to study the application of blind deconvolution to DBT data is proposed. Four different approaches to blind deconvolution are presented. Using the method described in Refs. [36] and [37] and noise regularization based on minimization of TV [33], a first estimate for the PSF is obtained. In addition, since the out-of-plane artifact is caused by features of high attenuation, representing sudden intensity variations, information about gradient magnitude is also taken into account during the deconvolution. The results are analyzed using real phantom data under 2D slice-by-slice visualization and 3D volume rendering with a compositing technique. The proposed methodology is also tested with one clinical DBT data set with a large calcification. To the best of our knowledge, this is the first study of out-of-plane

artifact reduction through this type of deconvolution and noise regularization in DBT data.

5.2. Materials and methods

5.2.1. Acquisition, reconstruction and visualization

To simulate high-density regions, a phantom built in our institution was used. A scheme and an image of the phantom are shown in Figure 5.1. The phantom contains two columns of aluminum disks embedded in an acrylic background (Figure 5.1 (b)). Considering the vertical direction, the aluminum disks were inserted in the central plaque of the phantom (dark gray in Figure 5.1 (a)). The disks have different diameters and 1 mm thickness. For this study, the first column schematically represented in Figure 5.1 (b) was considered (disks with diameters of 5.0 mm, 3.0 mm, 1.0 mm, 0.5 mm, 2.0 mm and 4.0 mm, from top to bottom).

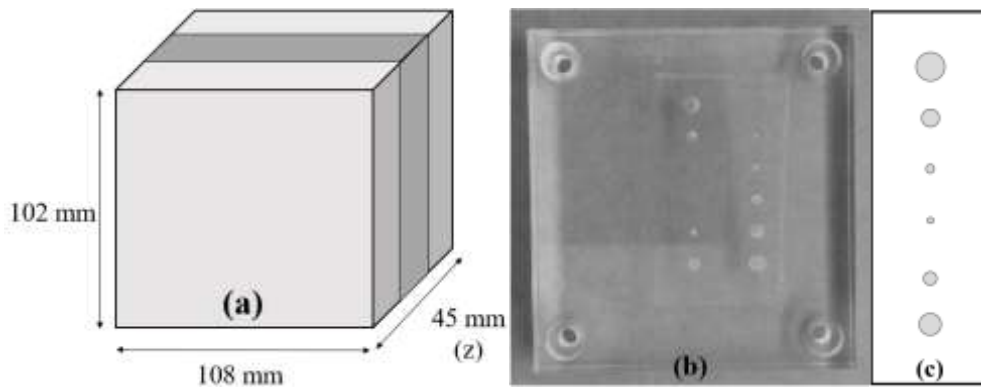


Figure 5.1. (a) Scheme of the acrylic phantom used. (b) Image of the phantom simulating breast tissue and lesions of high attenuation (aluminum disks of different diameters and 1 mm thickness). (c) Scheme of the disks in the first column (top to bottom): 5.0 mm, 3.0 mm, 1.0 mm, 0.5 mm, 2.0 mm and 4.0 mm, respectively.

Additionally, one clinical DBT data set with a large calcification of an anonymous patient was selected from a clinical facility database (Hospital da Luz S.A., Lisbon, Portugal). Both phantom and clinical data set were acquired with a Siemens MAMMOMAT Inspiration system (Siemens AG, Healthcare Sector, Erlangen, Germany). The reconstructions have voxel sizes of $0.085 \times 0.085 \times 1 \text{ mm}^3$ and were obtained with the manufacturer algorithm which uses Filtered Back Projection with some post-processing to reduce artifacts and image blurring [38].

The 3D visualization software was developed in C++ using the Visualization Toolkit library (VTK) version 7.1.0 [39, 40]. The methodologies in study were analyzed with 3D volume rendering visualization using ray casting and compositing. Different rendering parameters yield

different images. In this way, the parameters were fixed for all situations so that a correct comparison could be made.

5.2.2. PSF estimation in z-direction

The phantom was reconstructed in the xy plane and z-depth. As the out-of-plane artifact is present in the z-direction, its study was performed considering a PSF in the xz plane. By combining the methods described in Refs. [36] and [37], one xz plane containing the entire 0.5 mm disk distribution was chosen. In that plane, a region of interest (ROI) containing the 0.5 mm disk and some background (BG) was extracted. To reduce the noise variation, a TV minimization filter was applied to this region [33]. In order to achieve the minimum value of TV, several values of Lagrange parameter were applied to the longitudinal direction. From this fine analysis, Lagrange parameter = 156 has shown to be the one that allows the maximum reduction of TV and, for this reason, it was the chosen value for the application of the filter in this direction. In addition, the mean intensity of two BG ROIs was calculated and subtracted from that region. This region was thus considered as the first estimation of the PSF in z-direction.

5.2.3. Blind deconvolution

With blind deconvolution, an approximation of the true data u (deblurred) and the PSF can be recovered using observed data f (blurred) and an initial estimation of the PSF. As described in Figure 5.2, four approaches using blind deconvolution were studied in this work.

Pseudocode of the four iterative blind deconvolution algorithms	
1:	Open original data (xy plane, z depth) and rotate to xz plane, y depth: Observed Data (xz)
2:	Compute initial PSF estimate from the 0.5 mm disk: PSF
3:	Compute gradient magnitude mask: G_{mag}
If (Method 1)	
4:	Compute deblurred data with <i>deconvblind</i> over 10 iterations giving as input: Observed Data (xz), PSF
Elseif (Method 2)	
5:	Compute PSF estimate after 20 iterations (PSF_{refined}) with <i>deconvblind</i> giving as input: Observed Data (xz), PSF
6:	Compute deblurred data with <i>deconvblind</i> over 10 iterations giving as input: Observed Data (xz), PSF_{refined}
Elseif (Method 3)	
7:	Compute deblurred data with <i>deconvblind</i> over 10 iterations giving as input: Observed Data (xz), PSF, G_{mag}
Elseif (Method 4)	
8:	Compute PSF estimate after 20 iterations (PSF_{refinedmag}) with <i>deconvblind</i> giving as input: Observed Data (xz), PSF, G_{mag}
9:	Compute deblurred data with <i>deconvblind</i> over 10 iterations giving as input: Observed Data (xz), PSF_{refinedmag}
End	

Figure 5.2. Pseudocode of the four methods implemented in order to study the application of blind deconvolution in the reduction of out-of-plane artifacts in DBT data.

Method 1 is the simplest one. An initial estimation of the PSF, obtained as described in section 5.2.2, and rotated data (so that the deconvolution occurs in xz planes) were given as input to blind deconvolution. After a certain number of iterations, the deblurred data were obtained. The second method is similar to the first one, except that prior to data deblurring, the deconvolution algorithm ran for 20 iterations and the resulting PSF at the end of these iterations was given as input, together with the blurred data, for the deconvolution of the data itself. The choice of using a PSF after 20 iterations has relied on empirical studies made during the experiment.

Methods 3 and 4 are identical to methods 1 and 2, respectively, with the exception that as input, in addition to the PSF estimation and blurred data, a mask corresponding to the gradient magnitude of the same data was also considered. The gradient magnitude allows to detect sudden transitions of intensity corresponding to the high attenuation features which originate the artifacts. The use of this mask determines how much the voxel at the corresponding position in the input data is considered. In this way, these sudden changes were intended to have zero weight so that their influence on the deconvolution was minimal and to reduce the contrast-related image artifacts. After the gradient calculation, its complement was obtained and a binarization was carried out, where values of zero correspond to that transitions.

Using observed data as the reference, structural similarity index (SSIM) values calculated for

deblurred data at the end of each iteration were considered for stopping the iterative process. In order to preserve the fidelity of the results down to the smallest structure, SSIM was calculated over an area including the smallest disk (0.5 mm), considering all slices. The deconvolution algorithms were stopped when the SSIM dropped below 0.7.

The algorithms were implemented based on the MATLAB R2016b function *deconvblind* [41] and an Intel® Core™ i5-5200U CPU (2.20 GHz) @ 8 GB of memory computer was used.

5.2.4. Data analysis

The artifact in the z-direction was quantitatively analyzed through standard 2D slice-by-slice visualization and also through 3D volume rendering. For the 2D analysis, the artifact spread function (ASF) in the z-direction was calculated for each disk as in Eq. 5.2 in order to verify the consistency of the results at contrast and noise level.

$$ASF(z) = \frac{\mu_D(z) - \mu_{BG}(z)}{\sigma_{BG}(z)} / \frac{\mu_D(z_0) - \mu_{BG}(z_0)}{\sigma_{BG}(z_0)} \quad Eq. 5.2$$

In Eq. 5.2, z_0 is the slice location of the in-focus plane of the disk and z an off-focus plane. μ_D and μ_{BG} stand for mean pixel values in the ROI over the disk and BG, respectively, σ_{BG} stands for standard deviation in BG ROI. The values of μ_{BG} and σ_{BG} were obtained using two circular ROIs arranged on either side of the disk. The ASF curve is related with the extent of the artifact, namely its width. In this way, the full width at half maximum ($FWHM_{ASF}$) of a Gaussian curve fitted to the ASF of each disk, for each method, was calculated. To evaluate the accuracy between ASF points and the Gaussian curve, the coefficient of determination (R^2) was obtained for each fitting.

In volume rendering, when changing a camera's azimuth, its position rotates around the focal point. The result is a horizontal rotation (to the left or right) of the camera keeping the distance to the focal point constant [40]. In this case, two different angles were used to visualize the DBT data: 0° , which corresponds to the visualization parallel to the detector (i.e. projection is made on the xy planes along z); and 90° , which is the visualization perpendicular to the detector plate (i.e. projection is made on the xz planes along y). For the 3D quantitative evaluation using volume rendering, three figures of merit were obtained for each disk: FWHM (at 90°), noise (at 0° and 90°) and signal to noise ratio (SNR) (at 0° and 90°). The profile of each disk at 90° was obtained and the $FWHM_{90^\circ}$ of a Gaussian curve fitted to each profile was considered as an indicator of the spreading size of the artifact. In this case, the R^2 values was also obtained for

each fitting. Noise was calculated as in *Eq. 5.3* and, in order to study the relation between signal strength and BG noise, SNR was calculated using *Eq. 5.4*.

$$Noise = \frac{\mu_{BG}}{\sigma_{BG}} \quad Eq. 5.3$$

$$SNR = \frac{\mu_D}{\sigma_{BG}} \quad Eq. 5.4$$

For qualitative analysis, displays of the phantom and one clinical case obtained with volume rendering visualization at 0° and 90° are presented.

5.3. Results

The SSIM for deblurred data using observed data as reference remained above 0.7 until iteration number four (Figure 5.3). The SSIM values obtained for this iteration were 0.75, 0.71, 0.81 and 0.79 for method 1, 2, 3 and 4, respectively. In this way, the stop iteration was the fourth, achieved approximately after 3.7 minutes for the four methods.

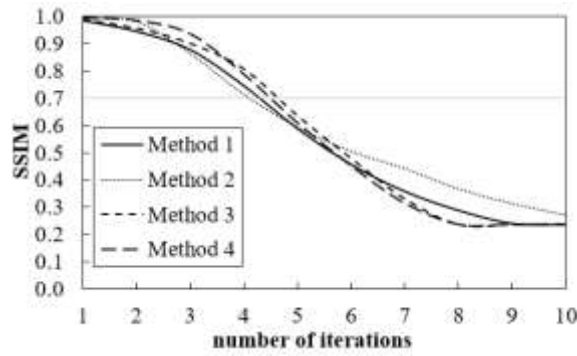


Figure 5.3. SSIM values obtained for deblurred data with each iterative blind deconvolution method, as a function of iteration number.

5.3.1. 2D analysis

The obtained ASF for deblurred data with each method and original data are presented from Figure 5.4 (a) to Figure 5.4 (f), for the six disks.

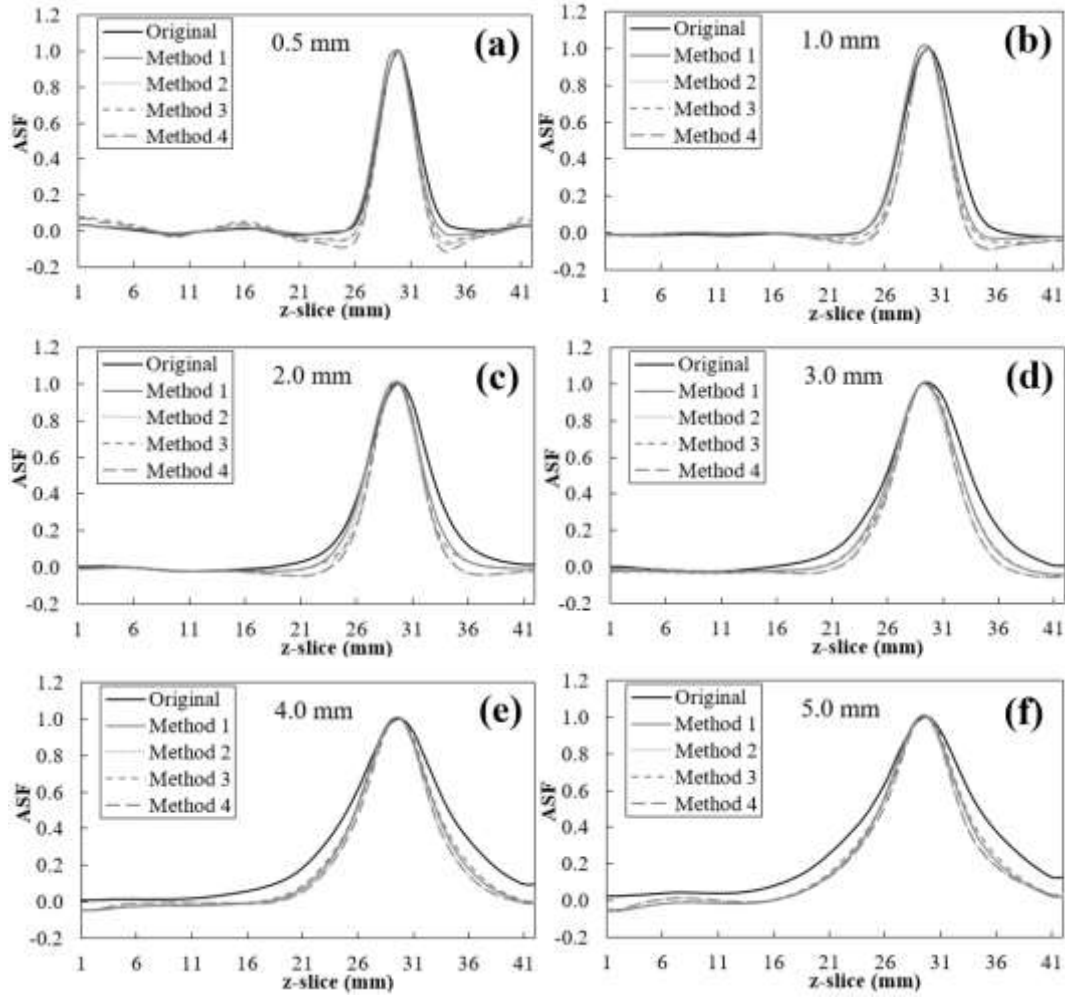


Figure 5.4. ASF curves obtained for each deconvolution method for the 0.5 mm (a), 1.0 mm (b), 2.0 mm (c), 3.0 mm (d), 4.0 mm (e) and 5.0 mm (f) disk.

The FWHM_{ASF} values estimated for each disk are given in Figure 5.5 and the R^2 values ranged between 0.96 and 1.0, revealing good adjustments. The variation, in percentage, between FWHM_{ASF} values of blurred and deblurred data with each method, are given in Table 5.1. The last row of Table 5.1 represents the average variation (in this case, reduction) of FWHM_{ASF} achieved with each method, considering the results in each disk.

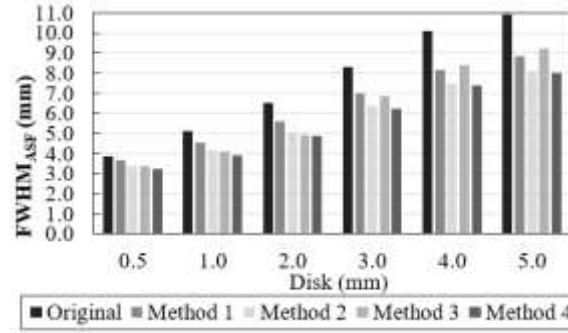


Figure 5.5. FWHM_{ASF} values of a Gaussian curve fitted to the ASF of each disk (0.5 mm, 1.0 mm, 2.0 mm, 3.0 mm, 4.0 mm and 5.0 mm), for each method and original data.

Table 5.1. Variation, in percentage, between the FWHM_{ASF} values achieved for the original data and after deconvolution with each method. Taking into account the different disks, the average of the obtained variations (reductions) with each method is presented in the last row.

$\Delta_{FWHM\ ASF} (\%)$				
Disk (mm)	Method 1	Method 2	Method 3	Method 4
0.5	-5.0	-12.5	-12.1	-16.0
1.0	-11.0	-19.0	-19.9	-23.1
2.0	-13.9	-21.8	-24.2	-25.2
3.0	-15.4	-23.2	-17.4	-25.2
4.0	-19.4	-25.7	-16.7	-26.8
5.0	-19.1	-25.5	-15.6	-26.4
mean	-14.0	-21.3	-17.7	-23.8

5.3.2. 3D analysis

The FWHM_{90°} values obtained for the profile of each disk at 90° are presented in Figure 5.6. Gaussian curve fittings for each profile showed R² values above 0.91.

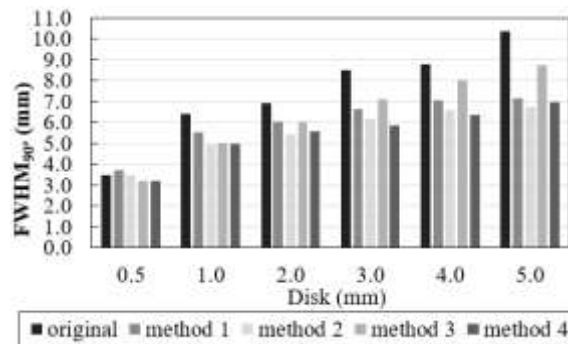


Figure 5.6. FWHM_{90°} values of a Gaussian curve fitted to the profile of each disk at 90° (0.5 mm, 1.0 mm, 2.0 mm, 3.0 mm, 4.0 mm and 5.0 mm), for each method and original data.

As for the FWHM_{ASF} values, the percentage of variation between FWHM_{90°} values of blurred and deblurred data with each method, are shown in Table 5.2.

Table 5.2. Variation, in percentage, between the FWHM_{90°} values achieved for the original data and after deconvolution with each method. Taking into account the different disks, the average of the obtained variations with each method is presented in the last row.

$\Delta_{FWHM\ 90^\circ} (\%)$

Disk (mm)	Method 1	Method 2	Method 3	Method 4
0.5	7.4	0.8	-7.4	-8.0
1.0	-13.5	-23.5	-22.1	-22.4
2.0	-12.9	-21.8	-12.5	-19.4
3.0	-21.5	-27.1	-16.1	-31.1
4.0	-19.7	-25.1	-8.4	-27.6
5.0	-30.9	-35.1	-15.6	-33.0
mean	-15.2	-21.9	-13.7	-23.6

Results of noise and SNR at 0° and 90° obtained in the phantom original and deblurred images for each disk are presented in Figure 5.7 and Figure 5.8, respectively. For a qualitative inspection, volume rendering images of the phantom at 0° and 90° are presented in Figure 5.9 and Figure 5.10, respectively.

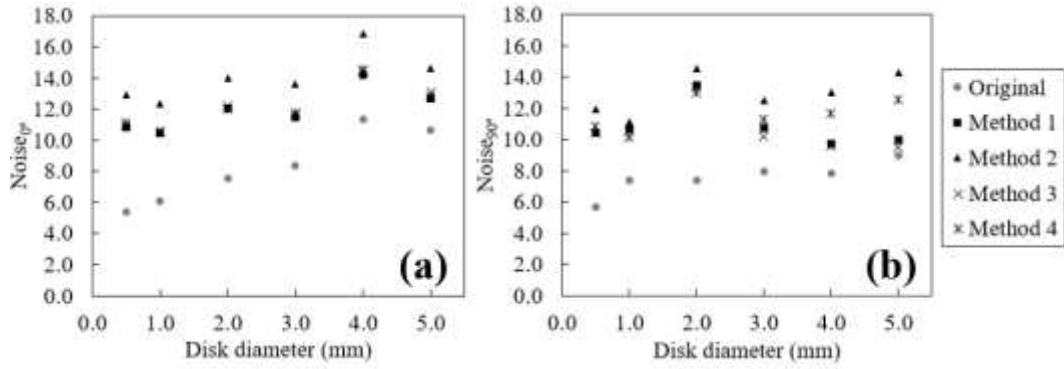


Figure 5.7. Noise at 0° (a) and 90° (b) obtained for phantom original and deblurred data with each method plotted as a function of disk diameter.

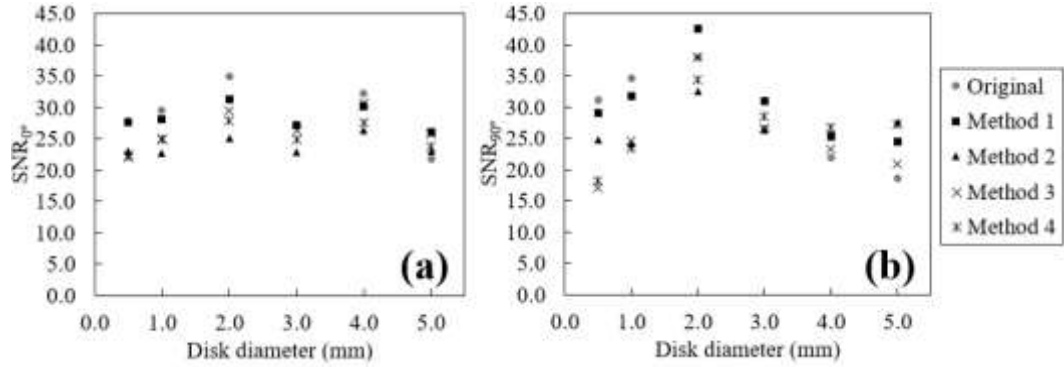


Figure 5.8. SNR at 0° (a) and 90° (b) obtained for phantom original and deblurred data with each method plotted as a function of disk diameter.

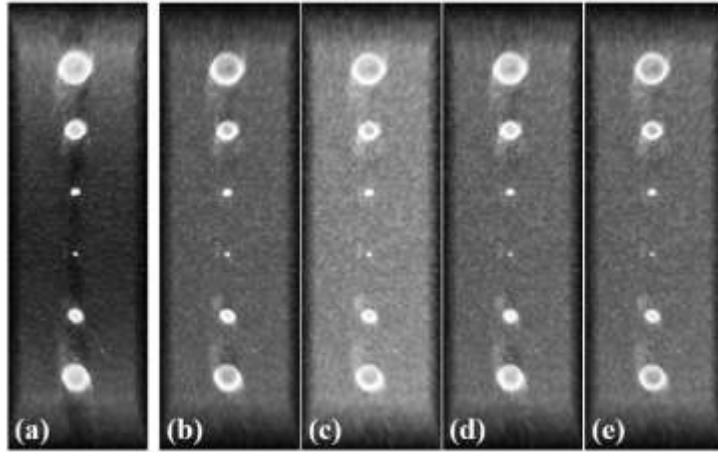


Figure 5.9. 2D displays of composite volume rendering visualization obtained at 0° for blurred (a) and deblurred data with method 1 (b), method 2 (c), method 3 (d) and method 4 (e).

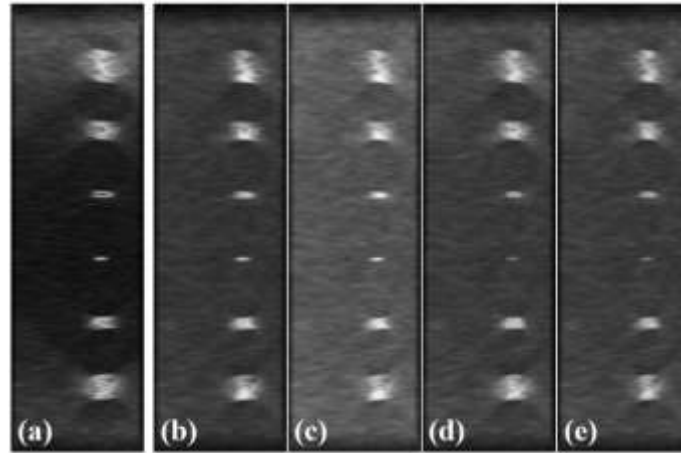


Figure 5.10. 2D displays of composite volume rendering visualization obtained at 90° for blurred (a) and deblurred data with method 1 (b), method 2 (c), method 3 (d) and method 4 (e).

To evaluate the consistency of the proposed methodology, method 4 was applied to one clinical data set and it was stopped at iteration number four. For clinical data, the method 4 took approximately 2 minutes to complete the four iterations. 2D displays of composite volume rendering obtained at 0° and 90° are shown in Figure 5.11.

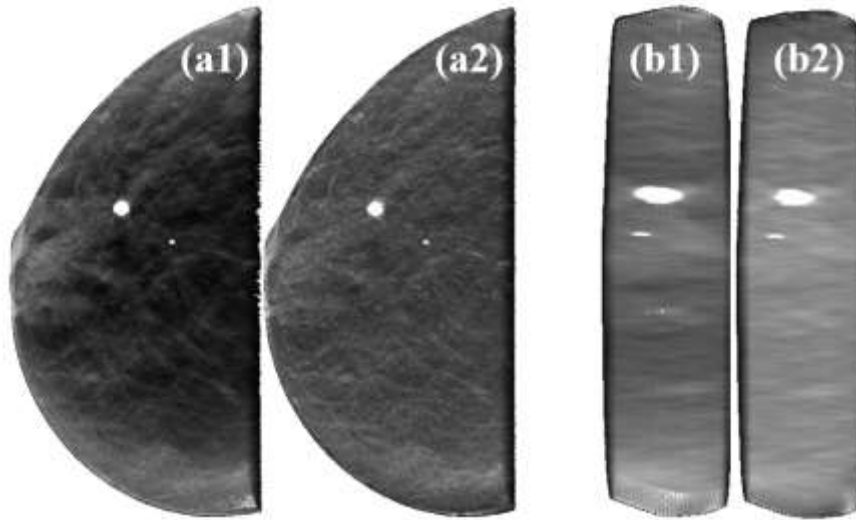


Figure 5.11. 2D displays of composite volume rendering visualization obtained at 0 and 90 degrees (a and b, respectively) for blurred (a1 and b1) and deblurred data with method 4 (a2 and b2).

5.4. Discussion

Four methods using TV minimization and blind deconvolution to reduce out-of-plane artifacts in DBT imaging have been presented. The artifact in the z-direction was quantitatively analyzed through standard 2D slice-by-slice visualization and with 3D volume rendering. The 2D evaluation was focused on the ASF curves of each disk and respective FWHMASF. The 3D analysis was done using the FWHM of disks profiles at 90° (FWHM90°), noise and SNR at 0° and 90°. 2D displays of volume rendering at 0° and 90° were considered for visual inspection.

In phantom data we observed that, regardless of the methodology and visualization approach, the objective of reducing the out-of-plane artifact was accomplished for all disk sizes (Table 5.1 and Table 5.2). In 2D, ASF was narrowed by all methods, for all disks, near the in-focus slice (slice number 30). Although all disks have a thickness of 1 mm (in z), it was clear that their diameter (in xy) affects the size of the artifact. For example, if we compare ASF of 1.0 mm disk (Figure 5.4 (b)) with ASF of 5.0 mm disk (Figure 5.4 (f)), it is clearly evident that for the larger diameter in xy, the artifact in the z-direction is larger.

These results were confirmed by FWHMASF values in Figure 5.5. The 5.0 mm disk presented the highest values of FWHMASF for all cases (original data and deblurred data) and the 0.5 mm disk the lowest FWHMASF values. Method 4 yielded the lowest FWHMASF values for all disks, resulting in the method that, on average, achieves a larger reduction on the ASF width: 23.8% (Table 5.1). With method 1 and 3, average reductions in FWHMASF of 14.0% and 17.7% were obtained, respectively, whereas method 2 achieved 21.3%. The R2 values obtained

after fitting a Gaussian curve to the ASF points were higher than 0.95, supporting the validity of the approximations. In this way, the FWHMASF values were considered as an estimate of the ASF curve width.

With volume rendering visualization, realistic computer-generated images of a 3D scene are obtained, yielding a true depth perception available through angular rotation. In this particular study, since the out-of-plane artifact propagates through multiple slices, we were interested in what happens in the z-direction. For that reason, a rotation of 90° was done and some quantitative analysis was performed on these rendered images. Through Figure 5.5 and Figure 5.6 we can see that the results from both visualizations followed the same pattern. Disks with largest dimensions at 0° (xy planes) have a higher FWHM value at 90° and also a higher deblurring after deconvolution methods. In Table 5.2, the 5.0 mm disk presented reductions in the spreading size of the artifact in the order of 30%, whereas 1.0 mm disk was at 20%. The mean values of the variations in Table 5.1 (last row) are comparable to those in Table 5.2. Taking all disks into account, the mean reductions of FWHM 90° values in volume rendering were 15.2%, 21.9%, 13.7% and 23.6% for method 1, 2, 3 and 4, respectively. The R2 values obtained after fitting a Gaussian curve to the volume rendering profiles at 90° were higher than 0.90, showing once again a good adjustment. Additionally, at 90° (Figure 5.10), there was a noticeable disk blur reduction in all deblurred images. This is in line with the numerical reduction of FWHMASF and FWHM 90° values. It is important to notice that, neither in the ASF nor in rendered images, were observed ringing artifacts often resultant from this type of deconvolution due to the limited bandwidth of the system. However, for the smaller disk, some overshoot values were observed in the ASF of methods 3 and 4 (although there are no negative effects in the quality of the images resulting from these methods). This can be explained by the location of the ROIs and by the fact that the ROI covering the 0.5 mm disk is the smallest one (being more sensitive to small variations). Additionally, PSF-based methods are known to cause edge artifacts which appear as an overshoot at the sharp transitions of the intensity of the phantom. To address this issue, an improvement in the PSF estimation should be considered.

Deconvolution methods introduced noise at 0° and 90° (Figure 5.7). The noise for each disk was measured considering a BG region surrounding the disk. In this way, the differences observed are essentially due to the spatial dependence of noise on DBT and not to the disk in question. Despite increasing it, deconvolution methods homogenized data noise level, with smaller differences obtained in the different spatial regions. These results are perceptible in Figure 5.9

and Figure 5.10. In original rendered images (Figure 5.9 (a) and Figure 5.10 (a)), it is visible that top and bottom phantom regions, which coincide with larger disks (5.0 mm and 4.0 mm, respectively), present a greater noise level than the center region (near the 0.5 mm disk). In phantom rendered images resultant from deconvolution methods (Figure 5.9 (b)-(e) and Figure 5.10 (b)-(e)), this difference between top/bottom regions and the center is not noticeable. On the other hand, SNR values of each disk at 0° were very close in the original and deblurred images (Figure 5.8 (a)). This means that although there has been an increase in noise level, the increase in signal strength on each disk has allowed a balance. SNR was preserved and the visualization of each disk was not affected. At 90° (Figure 5.8 (b)), a decrease in SNR values for the smaller disks (0.5 mm and 1.0 mm) was observed for all methods whereas, for the remaining disks, SNR values increased after deconvolution. By comparing the original (Figure 5.10 (a)) and deblurred phantom images (Figure 5.10 (b)-(e)), this decrease in SNR is expressed with a loss of contrast in smaller disks, in particular with methods 3 and 4. However, the diameter of the disks and their shapes after deconvolution remain intact, preserving the structural information correctly. In any case, this fact should be carefully analyzed in the future.

From numerical results of Table 5.1 and Table 5.2, we observed that giving as input a PSF which was already estimated after 20 iterations improves the results (method 2 and 4). This happens because the initial estimate of PSF (after 20 iterations) is closer to the true PSF, improving the deconvolution algorithm results for the same number of iterations. The difference between method 2 and 4 was the introduction of the gradient magnitude mask. Due to the high intensity of the disks, the deconvolution process can result in noise and blurring, essentially on smaller disks (such as 0.5 mm disk), because those are more susceptible to minor fluctuations. Results obtained with method 2 were an example of that. Considering Table 5.2, method 2 presented an equal or superior performance than method 4 on the largest disks (for example, 5.0 mm and 2.0 mm), whereas on a smaller disk (0.5 mm) the $FWHM_{90^\circ}$ even increased compared to the original data. The methods which used the information of the gradient magnitude mask, excluding the voxels around the disks during the restoration (assigning zero weight to them) were able to preserve this information, without blurring it. This resulted in improvements of 7.4% (method 3) and 8.0% (method 4) in the deblurring of the smaller disk.

The method which achieved the maximum mean reduction of $FWHMASF$ and $FWHM_{90^\circ}$ (method 4) was applied to the clinical data. The stop iteration according to the criterion of $SSIM \geq 0.7$ was the fifth. However, to keep the comparison with phantom data reliable, the clinical

images presented were obtained in iteration number four ($SSIM = 0.87$). The results obtained with the clinical data (Figure 5.11) were in accordance with the phantom results. There was a visible reduction of the big calcification blur at 90° , with an increase of noise level at 0° . The preliminary clinical data results here presented are intended to consolidate the results obtained with the phantom. Additional studies with volume rendering should be considered in the future. For example, the size of the calcification should be known in order to perform a quantitative study and thus to draw conclusions in this regard.

Although beam hardening artifact is extremely rare in DBT [42], it was observed in the images of the phantom at 0° (Figure 5.9), namely in the larger disks. This effect is present in the original image and remains evident in the images after deconvolution. In this way, it was not introduced by blind deconvolution itself. The phantom disks under study have a much higher density than the acrylic background, producing a very abrupt transition, thus resulting in this effect on the images. However, the images that radiologists analyze on daily basis are clinical images, similar to those in Figure 5.11, where this artifact was not observed, neither in the original data nor after the application of the blind deconvolution method.

For future work, we plan to improve the PSF estimation and test other deconvolution approaches, such as Wiener and Lucy-Richardson. Deconvolution should also be optimized for the noise level in data (not only at the moment of PSF estimation). In addition, as the presented methods may produce different gains and results with other scanners (with different angular range and number of projections), the study should be extended to different acquisition scenarios.

5.5. Conclusion

In DBT data acquisition there are two major aspects that greatly influence the reconstructed images: the low dose projections and the angular limit. These factors produce noisy images and out-of-plane artifacts where high intensity structures lay in the in-focus plane but also spread along other planes (which may lead to a significant decrease in DBT image quality in the areas surrounding a lesion of interest). As the out-of-plane artifact is a current drawback in DBT imaging, four methods based on TV regularization and blind deconvolution have been studied. The obtained results were analyzed through standard 2D slice-by-slice visualization and through 3D volume rendering.

The methodologies were tested with phantom and clinical DBT data. Both quantitative and

qualitative analysis performed showed the relevance of this approach in improving the image quality in DBT by reducing the out-of-plane artifact without introducing other artifacts, typical of the deconvolution in noisy data. This methodology can be very useful for future tools in DBT, such as computer assisted diagnosis and DBT- guided needle biopsies.

This study presents the particularity of visual analysis being performed through a truly 3D visualization (volume rendering). In this way, it is possible to get a sense of the impact of these algorithms on the data in a direct way, by visualizing the DBT data at once from several angles (in this case, from 0° and 90°).

References

1. Gennaro, G., et al., *Digital breast tomosynthesis versus digital mammography: a clinical performance study*. Eur Radiol, 2010. **20**(7): p. 1545-1553.
2. Brandt, K.R., et al., *Can Digital Breast Tomosynthesis Replace Conventional Diagnostic Mammography Views for Screening Recalls Without Calcifications? A Comparison Study in a Simulated Clinical Setting*. American Journal of Roentgenology, 2013. **200**(2): p. 291-298.
3. Bonafede, M.M., et al., *Value analysis of digital breast tomosynthesis for breast cancer screening in a commercially-insured US population*. ClinicoEconomics and Outcomes Research: CEOR, 2015. **7**: p. 53-63.
4. Gao, Y., et al., *Digital Breast Tomosynthesis Practice Patterns Following 2011 FDA Approval: A Survey of Breast Imaging Radiologists*. Acad Radiol, 2017. **24**(8): p. 947-953.
5. Destounis, S., A. Santacroce, and A. Arieno, *DBT as a Screening Tool and a Diagnostic Tool*. Current Breast Cancer Reports, 2017. **9**(4): p. 264-271.
6. Partyka, L., A.P. Lourenco, and M.B. Mainiero, *Detection of Mammographically Occult Architectural Distortion on Digital Breast Tomosynthesis Screening: Initial Clinical Experience*. American Journal of Roentgenology, 2014. **203**(1): p. 216-222.
7. Durand, M.A., et al., *Tomosynthesis-detected Architectural Distortion: Management Algorithm with Radiologic-Pathologic Correlation*. Radiographics, 2016. **36**(2): p. 311-321.
8. Sechopoulos, I., *A review of breast tomosynthesis. Part I. The image acquisition process*. Med Phys, 2013. **40**(1): p. 014301.
9. Good, W.F., et al., *Digital breast tomosynthesis: a pilot observer study*. AJR Am J Roentgenol, 2008. **190**(4): p. 865-9.
10. Gur, D., et al., *Digital breast tomosynthesis: observer performance study*. AJR Am J Roentgenol, 2009. **193**(2): p. 586-91.
11. Caumo, F., et al., *Digital Breast Tomosynthesis with Synthesized Two-Dimensional Images versus Full-Field Digital Mammography for Population Screening: Outcomes from the Verona Screening Program*. Radiology, 2018. **287**(1): p. 37-46.
12. Benedikt, R.A., et al., *Concurrent Computer-Aided Detection Improves Reading Time of Digital Breast Tomosynthesis and Maintains Interpretation Performance in a Multireader Multicase Study*. American Journal of Roentgenology, 2017. **210**(3): p. 685-694.
13. Balleyguier, C., et al., *Improving digital breast tomosynthesis reading time: A pilot multi-reader, multi-case study using concurrent Computer-Aided Detection (CAD)*. Eur J Radiol, 2017. **97**: p. 83-89.
14. Tchou, P.M., et al., *Interpretation Time of Computer-aided Detection at Screening Mammography*. Radiology, 2010. **257**(1): p. 40-46.
15. Suetens, P., *Medical image analysis*, in *Fundamentals of Medical Imaging* 2009, Cambridge University Press: New York. p. 159-189.

16. Venson, J.E., et al., *A Case-Based Study with Radiologists Performing Diagnosis Tasks in Virtual Reality*. Stud Health Technol Inform., 2017. **245**: p. 244-248.
17. Wu, T., R.H. Moore, and D.B. Kopans, *Voting strategy for artifact reduction in digital breast tomosynthesis*. Med Phys, 2006. **33**(7Part1): p. 2461-2471.
18. Abdurahman, S., et al. *Out-of-Plane Artifact Reduction in Tomosynthesis Based on Regression Modeling and Outlier Detection*. 2012. Berlin, Heidelberg: Springer Berlin Heidelberg.
19. Wicklein, J., et al. *Metal and calcification artifact reduction for digital breast tomosynthesis*. in *SPIE Medical Imaging*. 2017. SPIE.
20. Dustler, M., et al., *High-attenuation artifact reduction in breast tomosynthesis using a novel reconstruction algorithm*. Eur J Radiol, 2019. **116**: p. 21-26.
21. Kolitsi, Z., G. Panayiotakis, and N. Pallikarakis, *A method for selective removal of out-of-plane structures in digital tomosynthesis*. Med Phys., 1993. **20**(1): p. 47-50. doi: 10.1118/1.597060.
22. Sun, X., W. Land, and R. Samala. *Deblurring of tomosynthesis images using 3D anisotropic diffusion filtering*. in *Medical Imaging*. 2007. SPIE.
23. Bliznakova, K., Z. Bliznakov, and N. Pallikarakis. *An improved algorithm for out-of-plane artifacts removal in digital tomosynthesis reconstructions*. 2010. Berlin, Heidelberg: Springer Berlin Heidelberg.
24. Jerebko, A., *Out of plane artifact reduction in digital breast tomosynthesis and ct*, 2016, Siemens Healthcare GmbH.
25. Mall, S., et al., *The role of digital breast tomosynthesis in the breast assessment clinic: a review*. Journal of Medical Radiation Sciences, 2017. **64**(3): p. 203-211.
26. Jähne, B., *Digital Image Processing: Concepts, Algorithms, and Scientific Applications*. 3rd ed1995, Berlin, Germany: Springer.
27. Gonzalez, R.C. and R.E. Woods, *Digital Image Processing Using MATLAB*. 3rd ed2008, Upper Saddle River, New Jersey: Pearson Prentice Hall.
28. Ayers, G.R. and J.C. Dainty, *Iterative blind deconvolution method and its applications*. Optics Letters, 1988. **13**(7): p. 547-549.
29. Sidky, E.Y., et al., *Enhanced imaging of microcalcifications in digital breast tomosynthesis through improved image-reconstruction algorithms*. Med Phys, 2009. **36**(11): p. 4920-32.
30. Seyyedi, S., et al., *An object-oriented simulator for 3D digital breast tomosynthesis imaging system*. Comput Math Methods Med, 2013. **2013**: p. 250689-250689.
31. Michielsen, K., et al., *Design of a model observer to evaluate calcification detectability in breast tomosynthesis and application to smoothing prior optimization*. Med Phys, 2016. **43**(12): p. 6577-6587.
32. Krammer, J., et al., *Evaluation of a new image reconstruction method for digital breast tomosynthesis: effects on the visibility of breast lesions and breast density*. The British Journal of Radiology, 2019. **92**(1103).
33. Mota, A.M., et al., *Total variation minimization filter for DBT imaging*. Med Phys, 2015. **42**(6Part1): p. 2827-2836.
34. Mota, A.M., et al. *An iterative algorithm for Total Variation minimization in DBT imaging*. in *VipIMAGE 2015*, . 2015. Tenerife, Spain: CRC Press.
35. Mota, A.M., et al. *3D Total Variation Minimization Filter for Breast Tomosynthesis Imaging*. 2016. Malmö, Sweden: Springer International Publishing.
36. Hu, Y.-H., B. Zhao, and W. Zhao, *Image artifacts in digital breast tomosynthesis: Investigation of the effects of system geometry and reconstruction parameters using a linear system approach*. Med Phys, 2008. **35**(12): p. 5242-5252.
37. Richard, S. and E. Samei. *3D task-based performance assessment metrics for optimization of performance and dose in breast tomosynthesis*. in *SPIE Medical Imaging*. 2011. SPIE.
38. Siemens. *MAMMOMAT Inspiration - Tomosynthesis Option*. 2015 [cited 2020 February]; Available from: https://www.accessdata.fda.gov/cdrh_docs/pdf14/P140011c.pdf.
39. VTK. *Visualization Toolkit - VTK*. 2020 [cited 2020 February]; Available from: <http://www.vtk.org/>.

40. Schroeder, W., K. Martin, and B. Lorensen, *The Visualization Toolkit: An Object-oriented Approach to 3D Graphics*. 4rd ed2006, USA: Kitware.
41. MathWorks. *MATLAB deconvblind function*. 2020 [cited 2020 January]; Available from: <https://www.mathworks.com/help/images/ref/deconvblind.html>.
42. Geiser, W.R., S.A. Einstein, and W.-T. Yang, *Artifacts in Digital Breast Tomosynthesis*. American Journal of Roentgenology, 2018. **211**(4): p. 926-932.

Automatic classification of simulated breast tomosynthesis whole images for the presence of microcalcification clusters using deep learning CNNs

6

Ana M. Mota^{a*}, Matthew J. Clarkson^b, Pedro Almeida^a and Nuno Matela^a, "Automatic classification of simulated breast tomosynthesis whole images for the presence of microcalcification clusters using deep learning CNNs". Submitted to *Medical Physics*; accepted under Major Revision.

Medical Physics is a peer-review journal which have an impact factor of **4.071** and is ranked in the **Q1** in the domain: "Radiology, Nuclear medicine & Medical imaging".

^aFaculdade de Ciências, Instituto de Biofísica e Engenharia Biomédica, Universidade de Lisboa, Campo Grande 1749-016 Lisboa, Portugal;

^bDepartment of Medical Physics and Biomedical Engineering and the Centre for Medical Image Computing, University College London, London, UK.

* Corresponding author: ammota@fc.ul.pt

Abstract

Purpose: Microcalcification clusters (MCs) are one of the most important biomarkers for breast cancer, especially in cases of nonpalpable lesions. The vast majority of deep learning studies on Digital Breast Tomosynthesis (DBT) are focused on detecting and classifying lesions, especially soft tissue lesions, in small regions of interest previously selected by some other specific method. Only about 25% of the studies are specific for MCs and all of them are based on the classification of small pre-selected regions. A completely automatic and direct classification, which receives the entire image, without prior identification of any regions, is important and crucial for the usefulness of these techniques in a real clinical and screening environment. The main purpose of this work is to implement and evaluate the performance of Convolutional Neural Networks (CNNs) regarding to the classification of a complete DBT image for the presence or absence of MCs (without any prior identification of regions).

Methods: In this work, four popular deep CNNs are trained and compared with a new architecture proposed by us. The main task of these trainings was the classification of DBT cases by absence or presence of MCs. A public database of realistic simulated data was used and the whole DBT image was taken into account as input. DBT data were considered without and with pre-processing (to study the impact of noise reduction and contrast enhancement methods on the evaluation of MCs with CNNs). The area under the receiver operating characteristic curve (AUC) was used to evaluate the performance.

Results: Very promising results were achieved with a maximum AUC of 94.19% for the Googlenet. The second best AUC value was obtained with the new network implemented, CNN-a, with 91.17%. This CNN had the particularity of also being the fastest in terms of training and testing time, thus becoming a very interesting model to be considered in the future.

Conclusions: An automatic classification of DBT images based on deep CNNs for the presence or absence of MCs has been presented. Classifying the whole image according to the presence or absence of MCs is a difficult task due to the size of MCs and all the information present in an entire image. With this work encouraging outcomes were achieved in this regard, obtaining similar results to other studies for the detection of larger lesions such as masses. Moreover, given the difficulty of visualize the MCs, that are often spread over several slices, this work may have an important impact on the clinical analysis of DBT images.

Keywords: Digital breast tomosynthesis; microcalcifications; deep-learning; convolutional neural network; virtual clinical trial.

6.1. Introduction

Breast cancer is the most commonly diagnosed type of cancer worldwide [1]. Over the last three decades, mortality rates for breast cancer have dropped from its peak by 41%, likely reflecting advancements in treatment and earlier detection through increased screening programs [2]. However, in women, this disease is still the leading cause of cancer death [1].

Breast screening is crucial in identifying breast cancer at an early stage, when it can be better located and treated, thus reducing the breast cancer mortality. It is estimated that women who chose to participate in an organized breast cancer screening programme have 60% lower risk of dying from breast cancer within 10 years after diagnosis [3]. Until recently, these screenings and breast cancer detection in general were mainly performed using Digital Mammography (DM). However, as a result of its 2D nature, DM presents two major limitations: low sensitivity in dense breasts with pathology and low specificity due to normal tissue superposition [4].

The use of Digital Breast Tomosynthesis (DBT) has been confirming the potential of DBT to address these limitations. Initially, DBT was studied and approved in conjunction with DM, demonstrating an increase in breast cancer detection rates and a significant reduction in recall rates [4-9], particularly with dense breasts [6]. Currently, by including Synthetic Mammography (SM) generated from DBT data, DBT alone is approved as a stand-alone modality to replace DM [10-16].

One major drawback with DBT is its increase in interpretation time when compared to DM [17, 18]. Computer-Aided Detection (CAD) systems with DBT were implemented and evaluated in an attempt to shorten the reading time while maintaining the radiologist performance. In fact, some results are very encouraging with reading time reductions between 14% and 29.2% without loss of diagnostic performance [19-21].

On the other side, there are mixed observations about DBT technology for the detection of microcalcification clusters (MCs). Some studies have revealed inferior image quality for visibility of MCs with DBT [22-24] while others have not [25-27]. As MCs are one of the most important biomarkers for breast cancer [28, 29], especially in cases of nonpalpable lesions, another CAD approach that has been extensively studied with DBT is the use of these conventional CAD systems to assist in the correct detection of MCs [30-38]. However, despite the efforts and improvements already achieved (such as decreasing the false negative rate), due to the high false positive rates and low specificity, these CAD systems have not reached a level

of performance that can be translated into a true improvement in the real screening of breast cancer [39-42].

In recent years, the increase in computational power has allowed the development of deep learning artificial intelligence (AI) algorithms composed of multi-layered convolutional neural networks (CNNs). These AI systems have emerged as a potential solution in the field of automated breast cancer detection in DM and DBT [42]. In fact, recently, there have been several published large-scale studies where the aim was to analyze the performance of AI systems alone and also the performance of breast radiologists with and without AI [21, 43-50]. The AI systems under evaluation achieved a comparable or even improved cancer detection accuracy when compared with the human experts. With these promising results and the need for an automatic detection system for lesions in DBT and in screening, much research has been carried out in this regard. A brief summary of these studies is presented in Table 6.1.

Table 6.1. Summary of deep learning DBT studies (ROI: Region of interest, AUC: Area under the curve, pAUC: partial AUC).

Reference	Classification task	ROI / Patch / Image	model	Best metric
[51]	True MCs Vs. False positives	ROI (16x16)	Own	AUC: 0.93
[52]	Presence/absence of masses and architectural distortions	Patch (256x256)	Based on Alexnet	Accuracy: 0.8640
[53]	Presence/absence of masses	ROI (32x32x25)	Own	AUC: 0.847
[54]	True masses Vs. False positives	ROI (128x128)	Own	AUC: 0.90
[55]	True masses Vs. False positives	ROI (64x64)	Based on VGG16	AUC: 0.919
[56]	Positive (malignant, benign masses) Vs. Negative images	Image (224x224)	Based on Alexnet	AUC: 0.6632
[57]	Malignant Vs. benign masses	ROI (128x128)	Based on Alexnet	AUC: 0.90
[58]	Malignant Vs. benign masses	Image (256x256)	Own	AUC: 0.87

[59]	Presence/absence of MCs	Patch (29x29x9)	Based on [60]	pAUC: 0.880
[61]	Positive Vs. Negative volumes	Image (1024x1024)	Based on Alexnet, Resnet50, Xception	AUC: 0.854 (Alexnet)
[62]	Positive Vs. Negative volumes	Image (832x832)	Based on Alexnet, Resnet, Densenet and Squeezenet	AUC: 0.91 (Densenet)
[63]	Benign Vs. Malignant lesions	ROI (224x224)	Based on VGG19	AUC (MCs): 0.97
[64]	Positive Vs. Negative patches	Patch (512x512)	Based on Resnet	AUC: 0.847
[65]	Malignant Vs. benign masses Vs. normal	ROI (256x256)	Based on VGG16	AUC: 0.917, 0.951, 0.993 (malignant, benign, normal)
[66]	Malignant Vs. benign masses	ROI (224x224)	Based on Densenet121	AUC: 0.8703
[67]	BIRADS 0 Vs. BIRADS 1 Vs. BIRADS 2	Image (2200x1600)	Based on Resnet50	AUC: 0.912 (BIRADS 0 Vs. non-0)
[68]	Predict breast density	Image	Based on Resnet34	AUC: 0.952
[69]	True MCs Vs. False positives	ROI (128x128)	Based on Resnet18	AUC: 0.9765
[70]	Malignant Vs. benign Vs. normal images	Image (150x150)	Own	AUC: 0.89

The vast majority of these studies focus on detecting and classifying soft tissue lesions, such as masses [52-58, 65, 66]. Besides the fact that these are important lesions for the characterization of breast cancer, another reason for this is that, in this type of lesion, it is possible to greatly reduce the data input size through interpolation, without losing the spatial resolution required to observe the lesion (the same does not occur with MCs). In this way, faster transfer learning solutions, useful when there is a lack of available training data (as in the case of DBT), can be used with very positive results [54-57, 65, 66]. Even in cases where only regions of interest (ROIs) and not full images are selected, such resizing is usually carried out. Furthermore, the

vast majority of the works use ROIs or patches where objectively there is or is not a lesion [56, 58, 61, 62, 67, 68, 70], instead of using the whole image or volume. The use of the whole image or volume is important to contextualize the lesions but also to make the classification a useful and quick tool in screening, where an image/volume should ideally give some type of objective and direct outcome.

One of the biggest challenges involving DBT in AI is the lack of a large, properly labelled public database. All studies mentioned in the Table 6.1 use private databases, making generalization and a fair comparison between different studies impractical [71]. Recently, two publicly-accessible annotated DBT datasets that will facilitate the evaluation and validation of AI algorithms were released. Buda et al. made publicly available a large-scale dataset of DBT data. It contains 5610 studies from 5060 patients: 5129 are normal cases (no abnormal findings), 280 are cases where additional imaging was needed but no biopsy was performed, 112 are benign biopsied cases and 89 are cases with proven cancer. This dataset includes masses and architectural distortions and was used to train and test a single-phase deep learning detection model that reached a baseline sensitivity of 65% at 2 false positives per DBT volume [72]. The other dataset resulted from the advancement of in-silico tools. The Virtual Imaging Clinical Trial for Regulatory Evaluation (VICTRE) project was created for the evaluation of the imaging performance of DBT as a replacement to DM for breast cancer screening. In VICTRE, the whole imaging chain was simulated with state-of-the-art tools and a total of 2986 virtual realistic patients were generated and imaged with both modalities. The positive cohort (that comprises malignant spiculated masses and MCs) included 1944 and 1042 virtual patients with and without lesions, respectively [73].

In this paper, fully automatic methods based on deep learning were studied for classifying DBT data. The aim is to input a whole DBT image and have a direct answer about the absence or presence of MCs, without the need for prior identification of lesions in specific regions and thus completely automate the process of DBT classification. Four existing popular networks were considered and compared with a new network proposed by us for this purpose. In order to study the impact of some pre-processing methods in increasing the visibility of MCs, the input data was considered with and without pre-processing. The VICTRE's public database was used. To the best of our knowledge, this is the first study of automatic classification specifically dedicated to the presence or absence of MCs in whole DBT images.

6.2. Materials and methods

6.2.1. Database

This study was centered on the database created for the VICTRE trial [73]. Synthetic images of virtual patients were obtained using an in-silico version of the Siemens Mammomat Inspiration DBT system using Monte Carlo x-ray simulations. These data are available to the public in the Cancer Imaging Archives [74]. Physical compression of left breasts was considered in the craniocaudal (CC) orientation. In this database, the cases are divided into absence and presence of lesions and also according to the density of the breast (fatty, scattered, heterogeneous and dense). The absent cases have no findings and each case with lesions present contains: four spiked masses with a 5-mm nominal diameter and mass density 2% higher than normal glandular tissue and four MCs consisting of 5 calcified lesions modelled as 195, 179, and 171 μm of solid calcium oxalate. In this study we included cases without (“absent”) and with MCs (“presentMCs”).

Table 6.2 presents a detailed summary of the dataset selected for this work. The reconstructed cases have different dimension in x, y and z, depending on breast density: $1624 \times 1324 \times 62$, $1421 \times 1024 \times 57$, $1148 \times 753 \times 47$ and $1130 \times 477 \times 38$ for fatty, scattered, heterogeneous and dense breasts, respectively, with a voxel size of $0.085 \times 0.085 \times 1 \text{ mm}^3$. For the absent category, five slices proportionally spaced between the first and the last slice were selected for each case (for example, as fatty cases have 62 slices: slices 1, 17, 33 49 and 62 were selected; as dense cases have 38 slices: slices 1, 11, 21, 31 and 38 were chosen). On the other hand, for the presentMCs class, slices containing the center of the cluster were selected for each case (in some cases, two clusters had their center on the same slice). Numerically, we adopted the usual distribution of breast density in the population: 10% fatty, 40% scattered, 40% heterogeneous and 10% dense [75] and an approximate balance between cases without and with lesions.

Table 6.2. Detailed summary of the VICTRE data selected for this study.

Density	Absent		PresentMCs	
	Nr. of cases	Nr. of slices	Nr. of cases	Nr. of slices
Fatty	20	100	25	99
Scattered	80	400	100	386
Heterogeneous	80	400	100	371
Dense	20	100	25	93
TOTAL		1000		949

6.2.2. Data pre-processing

In VICTRE's database, the reconstructed data have signal contamination outside the breast region, i.e., in the background (BG). This information is worthless for training the networks and, when present, slows down the process as pixels without any useful information end up contributing to the mathematical operations involved. In this way, through binarization and region growing operations, binary masks that keep information belonging to the breast and make everything else zeros were created ("BG suppression"). This step was applied to the original data and after all the other types of processing.

The very low dose projections acquired within a limited angular range in a DBT examination results in low statistics (high noise level) in the reconstructed images and data insufficiency. For this reason, image denoising methods are very important in order to improve the image quality of DBT data. Total variation (TV) minimization algorithms have attracted considerable attention in the field because of its ability for smoothing images while preserving the edges. Studies applying TV minimization to DBT data have shown excellent results so far [75-79]. This methodology was applied during pre-processing step. Minimization of TV greatly improves the contrast-to-noise ratio by reducing the noise. In this way, in order to also increase the contrast, two other techniques were studied. The contrast-limited adaptive histogram equalization (CLAHE) technique was implemented to increase the contrast of all breast structures in general and a simpler operation was applied to increase the contrast of structures with greater intensity, such as MCs, in particular. Since we wanted to study whether image noise reduction and/or contrast have any impact on CNNs training, some combinations of these methods were made, resulting in six different pre-processing approaches (Figure 6.1) and the next paragraphs explain how these operations were carried out.

Pre-processing 1:	noise reduction	→	BG supression	
Pre-processing 2:	contrast enhancement	→	BG supression	
Pre-processing 3:	noise reduction	→	contrast enhancement	→ BG supression
Pre-processing 4:	contrast enhancement	→	noise reduction	→ BG supression
Pre-processing 5:	normData ²	→	BG supression	
Pre-processing 6:	normData ²	→	noise reduction	→ BG supression

Figure 6.1. The six pre-processing methodologies implemented in order to reduce noise and amplify the visibility of the MCs (BG: background, normData: Data normalized between 0 and 1).

Pre-processing 1: As DBT data is composed by a high level of noise resulting from the acquisition of low dose projections, the application of a noise reduction filter was analyzed.

This filter consists in minimizing the TV of the data, allowing to significantly reduce the noise while preserving the edges and lesion resolution (which is a very important factor when the structures under analysis are small MCs). TV is a measure of pixel intensity variation in an image and increases significantly in the presence of noise. In each pre-processing that included this filter, several Lagrange multipliers were tested to study which allowed the minimum TV value [77] and 14 was the chosen value for the application of the filter in all cases.

Pre-processing 2: The CLAHE technique [80] was implemented based on the MATLAB R2020a function `adapthisteq` [81] to enhance the contrast of the images and the MCs. With this technique, the contrast in homogeneous areas is limited to avoid the amplification of noise. The contrast transformation function is calculated in small regions of the image individually, rather than in the whole image and neighboring regions are then combined through bilinear interpolation to eliminate artificially induced boundaries.

Pre-processing 3 and 4: The techniques described for pre-processing 1 and 2 were combined and used together by varying the order by which each one is applied. These steps (3 and 4) were also included since techniques 1 and 2 could complement each other and, through preliminary studies, it was possible to conclude that their order of implementation show differences in the appearance of the final image. In pre-processing 3, the TV minimization filter for noise reduction was first applied, followed by the contrast enhancement technique. For pre-processing 4, the application was in the opposite order, first contrast enhancement technique and then noise reduction.

Pre-processing 5: The data intensity was first normalized between 0 and 1 and then squared to attenuate the lower values, highlighting the higher ones belonging to the MCs. With this filter our aim was to specifically increase the contrast of regions of higher intensities.

Pre-processing 6: The method applied in pre-processing 5 was followed by the TV minimization filter, as described in pre-processing 1.

In order to homogenize the data, as well as to find a balance between training time/memory and the necessary spatial resolution for the visibility and conspicuity of MCs, all data were resized in x and y to 512×512 . No adjustments have been made in the z direction since training is performed slice-by-slice. The images were converted into TIFF slices of 8-bit and input data were normalized using the *zerocenter* method.

6.2.3. CNNs

Since it was crucial to maintain image spatial resolution under certain limits to allow the detection of the small MCs, it was not possible to reduce the image dimension to values such as 224×224 or 227×227 , which are the most used in pretrained networks for transfer learning. Our approach was then to train from scratch four architectures that already exist: Alexnet [82], Googlenet [83], Resnet18 [84] and Squeezenet [85]. The choice of these popular pre-trained networks was based on the comparison of each model speed and accuracy [86].

In addition, to alleviate some computational efforts, one faster and lighter new architecture, based on Alexnet, is proposed by us: CNN-a (Figure 6.2).

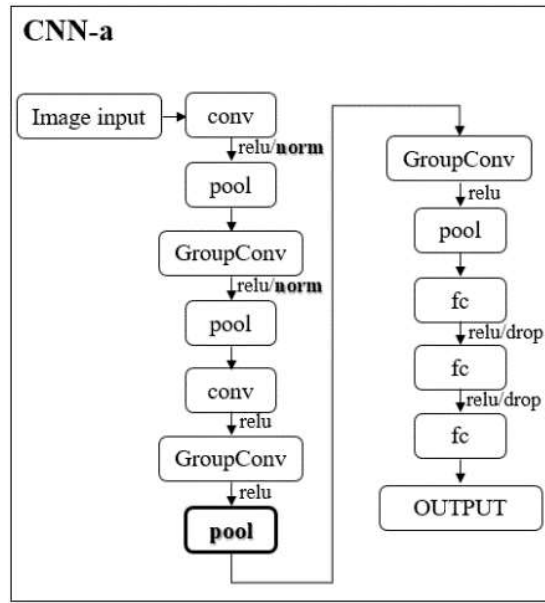


Figure 6.2. Illustration of CNN-a that resulted from the modifications made (bold) to the Alexnet architecture. *Conv* and *GroupConv*: convolutional and grouped convolutional layers, respectively; *pool*: max pooling layers; *fc*: fully connected layer; *relu*: rectified linear unit layer; *norm*: batch normalization layer; *drop*: dropout layer.

In CNN-a, the channel-wise local response normalization layers were replaced by batch normalization layers (“norm”) and a new max pooling layer with stride 2, padding 0 and size 3x3 was added between the two grouped convolutional layers. These modifications are the result of several empirical trial and error studies conducted by us during the experiment.

6.2.4. Methodology pipeline

Figure 6.3 shows the pipeline followed in this work. Absent and presentMCs data samples were selected and the described pre-processings were applied. The training dataset was used to train

the CNNs from scratch and the testing dataset was used after training, to evaluate the performance of the trained CNNs.

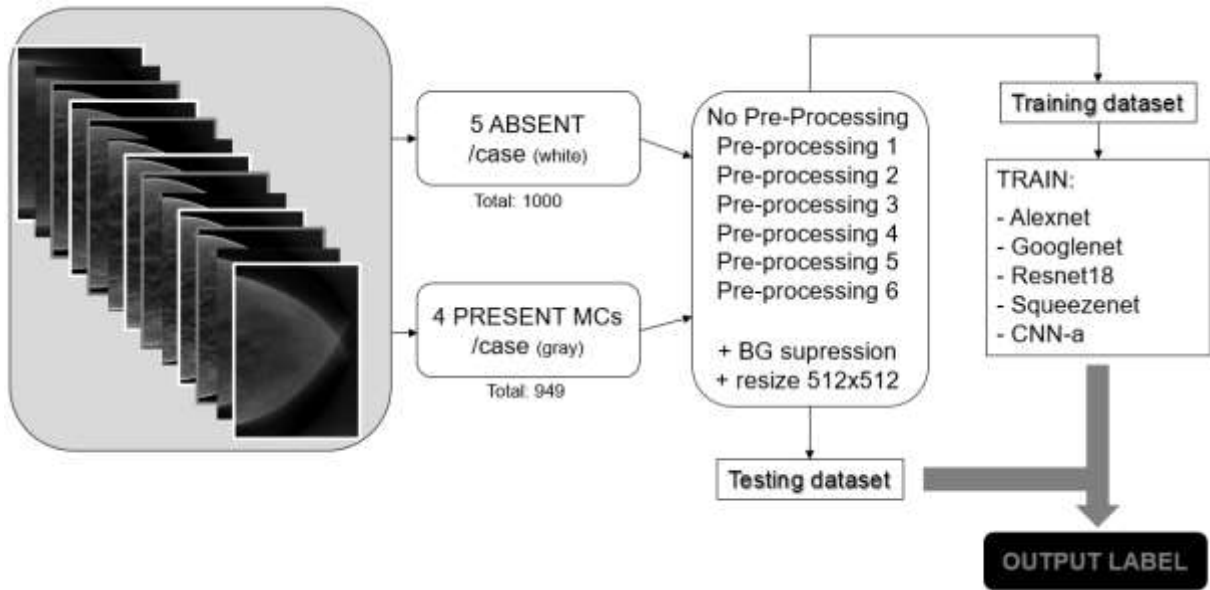


Figure 6.3. Summary of the methodological pipeline followed in this work.

6.2.5. Training options

The k-fold technique was used as the cross-validation method to estimate the generalization error of the learning process. The data set used was divided into $k=3$ subsets, i.e., each network was trained and tested three times with different data sets, always according to the proportion of $2/3$ of the cases for training and $1/3$ for testing. Since the split was performed at the patient level, all the images of the same patient were in either the training set or the test set. Training data augmentation was used through random reflection in the left-right direction (to simulate the inclusion of examples of right breasts) and data rotation between ± 20 degrees.

The CNNs were trained using the stochastic gradient descent optimizer with momentum 0.9 to minimize the cross-entropy loss for classification. The maximum number of epochs was 200 with a mini-batch size of 32 and a learning rate of 1×10^{-3} . Besides the 3-fold cross validation, an L2 regularization term of 5×10^{-3} was introduced in the loss function to prevent overfitting.

6.2.6. Evaluation metrics

Classification problems usually involve distinguishing between two classes. In the case of medical imaging, this distinction is usually made between the absence or presence of abnormalities or between benign/malignant lesions. In our work, the objective was to

distinguish between the absence or presence of MCs. Sensitivity, specificity, accuracy and area under the receiver operating characteristic (ROC) curve (AUC) were considered to evaluate the performance. The analysis of only the first three metrics can be limitative because they depend on the defined threshold to accept a case as presentMCs or absent. In this way, we used the AUC (positive class: presentMCs) as a summary tool that contains the space of all these possible thresholds.

Differences in performance of each classifier were tested using a statistical t-test. A two-tailed $p\text{-value} < 0.05$ was considered to indicate a significant difference.

6.3. Results

6.3.1. Data pre-processing

All the steps involved in the BG suppression are presented through an example case in. The original data was first binarized (Figure 6.4 (b)), the holes in the image were filled (Figure 6.4 (c)), the largest resultant object was selected (Figure 6.4 (d)) and the complete binary mask was achieved by performing region growing in (Figure 6.4 (e)). The profile traced for the white ROI (lower right corner of (a) and (f)) shows the cleaning effect of it.

This methodology was included in all pre-processing approaches, as mentioned in section 6.2.2. Zooming in on one MC (Figure 6.5), we can see the different results achieved in this type of lesions with each pre-processing method.

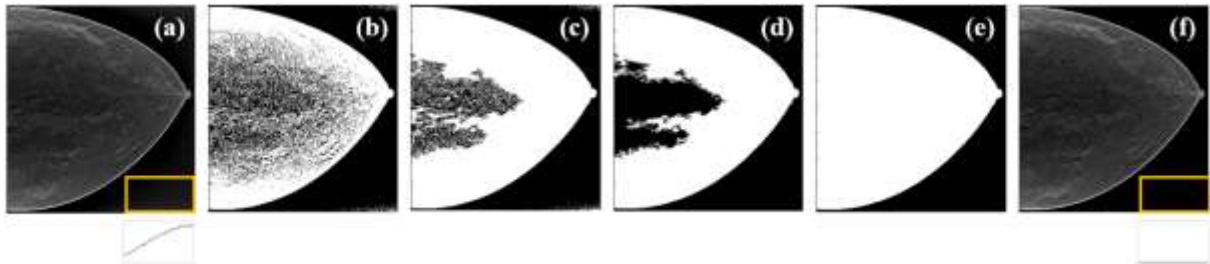


Figure 6.4. a) Data with contaminated BG; b) First binary image; c) Filled binary image; d) Largest object extracted from binary image; e) Result from region growing; f) Final image with BG corrected after binary mask from e) applied to a).

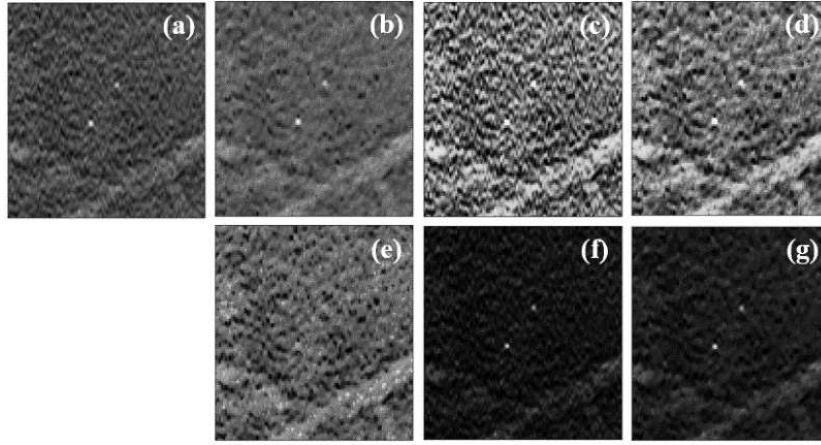


Figure 6.5. a) Original data without pre-processing; b) Pre-processing 1 (minimization of TV); c) Pre-processing 2 (CLAHE); d) Pre-processing 3 (minTV + CLAHE); e) Pre-processing 4 (CLAHE + minTV); f) Pre-processing 5 (dataNorm2); g) Pre-processing 6 (dataNorm2 + minTV).

6.3.2. Performance analysis

Our research was guided by the AUC results obtained for the different architectures and pre-processing methods. As mentioned above, the training and testing were repeated three times (3-fold cross validation) using three distinct datasets. The averaged performances and standard deviation values found over the three folds are shown in Table 6.3.

Table 6.3. Performance results of CNNs trained with original data and with data resulting from the pre-processing methodologies, in terms of mean AUC.

	AUC (%): mean \pm stdDev				
	Alexnet	Googlenet	Resnet18	Squeezenet	CNN-a
Original data	87.92 \pm 2.01	90.14 \pm 0.38	86.84 \pm 2.62	87.43 \pm 0.78	89.79 \pm 1.23
Pre-processing 1	87.35 \pm 1.63	88.38 \pm 1.12	87.96 \pm 0.96	88.78 \pm 0.99	90.66 \pm 0.15
Pre-processing 2	87.29 \pm 0.78	93.02 \pm 3.59	86.42 \pm 3.26	86.84 \pm 3.82	86.95 \pm 0.97
Pre-processing 3	88.61 \pm 0.43	94.19 \pm 1.12	86.33 \pm 1.46	82.15 \pm 1.51	85.80 \pm 1.73
Pre-processing 4	90.82 \pm 1.29	94.15 \pm 1.54	90.13 \pm 0.32	86.33 \pm 6.31	89.07 \pm 1.62
Pre-processing 5	87.62 \pm 0.35	88.65 \pm 4.27	90.44 \pm 0.41	85.18 \pm 2.78	89.54 \pm 2.63
Pre-processing 6	87.47 \pm 1.13	89.76 \pm 1.76	89.00 \pm 1.33	84.09 \pm 3.13	91.17 \pm 0.07

In Table 6.4 are the p-values calculated to study the measurable statistical differences between the best mean AUCs obtained in Table 6.3.

Table 6.4. Levels of significance (p-values) obtained from the statistical analysis of the difference between the best mean AUCs found.

p-value	Googlenet preProc3 (94.19 ± 1.12)	Resnet18 preProc5 (90.44 ± 0.41)	Squeezenet preProc1 (88.78 ± 0.99)	CNN-a preProc6 (91.17 ± 0.07)
Alexnet preProc4 (90.82 ± 1.29)	0.027 (Alexnet < Googlenet)	0.654	0.095	0.662
Googlenet preProc3 (94.19 ± 1.12)		0.006 (Googlenet > Resnet18)	0.003 (Googlenet > Squeezenet)	0.010 (Googlenet > CNN-a)
Resnet18 preProc5 (90.44 ± 0.41)			0.055	0.038 (Resnet18 < CNN-a)
Squeezenet preProc1 (88.78 ± 0.99)				0.014 (Squeezenet < CNN-a)
For p-values < 0.05 (in bold) there is a significant difference				

Considering only the best results obtained for averaged AUC, Figure 6.6 shows the ROC curves of CNN network trained with the respective data. These curves were obtained by averaging between the ROC curves of each fold. Additionally, Figure 6.7 analyzes the values of the respective sensitivities, specificities and accuracy in detecting the cases with MCs.

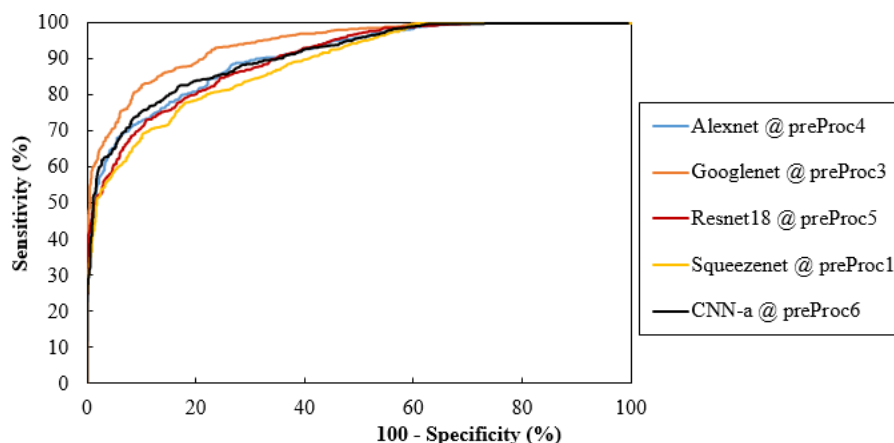


Figure 6.6. Comparisons of ROC curves for the CNNs and training data with the best AUC values.

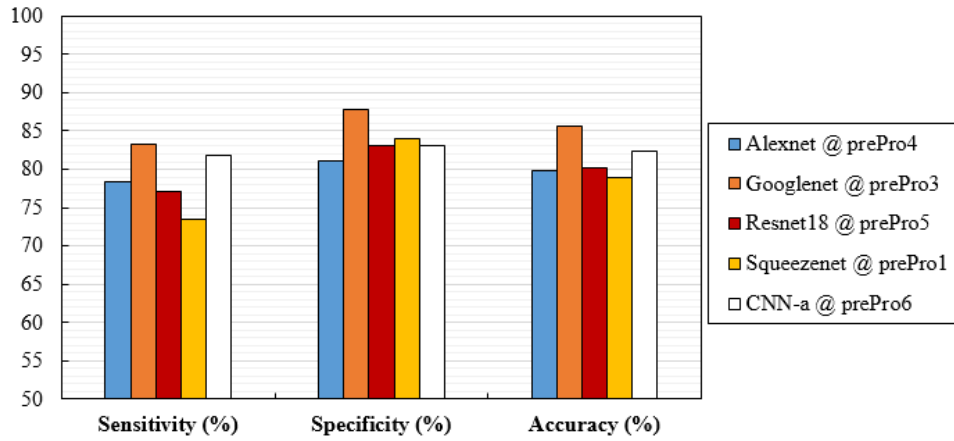


Figure 6.7. Values of sensitivity, specificity and accuracy obtained with the architectures trained with pre-processed data that achieved the best mean AUC.

6.3.3. Density influence

Breast density interferes with the detection of lesions [87]. In this way, it was important to explore the influence of density on the specific detection of MCs with these CNNs trained by these datasets. For this purpose, the training dataset were not changed, i.e., the CNNs were trained including all breast densities, but they were tested separately with specific datasets for each breast density. The results, in form of AUC, are shown in Figure 6.8.

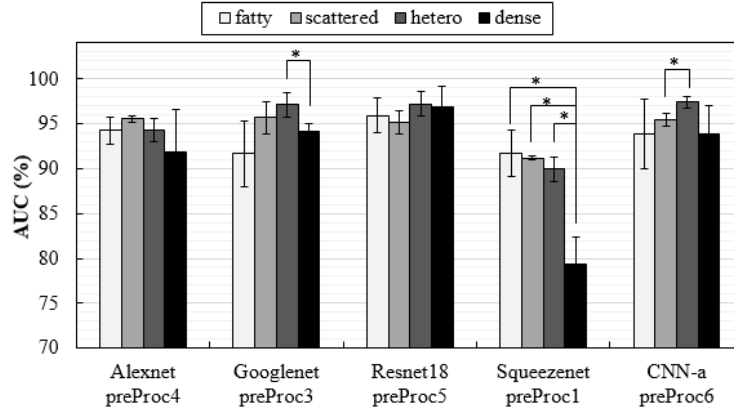


Figure 6.8. AUC values obtained with test datasets composed by the four different breast densities separately (* $p < 0.05$: there is a significant difference between groups).

The training that provided the best performance (Googlenet @ pre-processing 3) required a training time of approximately 9 h for all 3-folds (using a NVIDIA Quadro P4000 GPU). On the other hand, the fastest training and second best performance was obtained, simultaneously, for our CNN-a with data from pre-processing 6. Table 6.5 shows the training and inference times for all CNNs.

Table 6.5. Training times, in hours, needed for each CNN after 3-fold cross validation and mean inference time (in seconds) needed to classify each image.

	Training time (hours)	Inference time / slice (secs)
CNN-a	2.4	0.0057
Alexnet	4.1	0.0062
Squeezenet	4.4	0.0083
Resnet18	7.8	0.0143
Googlenet	8.9	0.0158

6.4. Discussion

In this work, the training from scratch of four popular CNNs and a new architecture proposed by us was investigated. Given as input the whole DBT image (and not only some specific ROIs), the classification of cases by absence or presence of MCs was the main task of these trainings. Original data and data resulting from pre-processing methods (to increase MCs visibility) were considered. The DBT dataset used for training and testing are from the publicly database available at The Cancer Imaging Archive website [74].

In order to avoid useless complex mathematical operations, all the information outside the breast region was eliminated. In four steps, an automatic methodology that creates a binary image where only the information inside the breast is considered was implemented. The comparison between the contaminated data and the data with complete suppression of BG signal can be observed through the profiles of the yellow regions in Figure 6.4 (a) and Figure 6.4 (f), respectively. This operation represented a difference of about 5% in training times, without performance losses, and it is usually done in this type of CNNs training.

Data pre-processing can be very useful when training CNNs from scratch to facilitate the detection and classification processes. In this work, both original data and data resulting from different pre-processing methods were considered as input. A comprehensive study of different methods to make the MCs more visible to the algorithms was carried out.

In original data, the MCs showed reasonable contrast to the naked eye (Figure 6.5 (a)). This highlight can be compromised due to their size and the presence of noise and other structures that can make them less visible. Both pre-processing 1 and pre-processing 2, had a great influence on MCs data. Pre-processing 1 smoothed the region around the MCs, preserving its edges (Figure 6.5 (b)), while pre-processing 2 contributed to an increase in contrast between all

structures, whether they were MCs or not (Figure 6.5 (c)). We thought it might be interesting to combine a technique that is essentially for noise reduction (TV minimization) with a CLAHE technique and, in this way, pre-processing 3 and 4 corresponding to Figure 6.5 (d) and Figure 6.5 (e), respectively, were implemented. While, visually, the MCs stand out from the surrounding noise in Figure 6.5 (d), in Figure 6.5 (e) where the contrast enhancement was applied first and the noise reduction latter, the MCs appear to fade. Additionally, for its simplicity, another method based on squared normalized data was also studied (pre-processing 5). This operation worked quite well when it comes to highlighting high intensity structures (Figure 6.5 (f)). The application of the TV minimization filter to these data (pre-processing 6) also resulted in a reduction of anatomical noise that allows for greater differentiation of the MCs, as can be seen in Figure 6.5 (g).

This descriptive analysis is in line with the numerical results obtained for the trained CNNs. From Table 6.3, it can be seen that not only the type of input data have affected the results, but also the CNN architecture itself. In fact, the best AUC value of each CNN was achieved with different input data. Googlenet showed the best AUC with data processed with method 3 (94.19%), CNN-a with method 6 (91.17%), Alexnet with method 4 (90.82%), Resnet18 with method 5 (90.44%) and Squeezenet with method 1 (88.78%). CNNs trained with original data didn't generate a maximum of AUC. However, all the AUC values were higher than 86%, showing that, even without any pre-processing, this could be an option. As shown in Figure 6.9 (a), for cases where the MCs are in a region with less noise and are more evident, all the CNNs get a correct classification in the original data. On the other hand, despite the efforts to reduce noise and increase contrast, some cases like the one in Figure 6.9 (b) were incorrectly classified as negative by all CNNs, even varying the pre-processing. Although pre-processing 2 didn't contribute to a maximum either, it resulted in the third best AUC for Googlenet. From Table 6.3, it is also possible to conclude that Googlenet was the most sensitive CNN to data contrast since its best results of AUC were obtained with methods where the contrast enhancement operation was performed. In the example of a case where MCs is in a region with other structures also of greater contrast (Figure 6.9 (c) and (d)), Googlenet took advantage of pre-processing 3 and was the only CNN to correctly classify this case. As a matter of fact, the Googlenet trained with data processed by method 3 presented significantly higher values in the detection of cases with MCs (p-value < 0.05, Table 6.4). This superiority is quite visible in the isolated ROC curve in the Figure 6.5. The second-best performance corresponded to the CNN-

a trained with data from pre-processing 6, being this superiority significant in relation to Resnet18 and Squeezenet Table 6.4. In Figure 6.9 (e) there is a case of a MCs that is masked and that was only detected by CNN-a after pre-processing 6 (Figure 6.9 (f)). Thus, and in agreement with the results in Table 6.3, we can assume that it is the combination of both factors (data type and CNN) that determines the result of a correct classification.

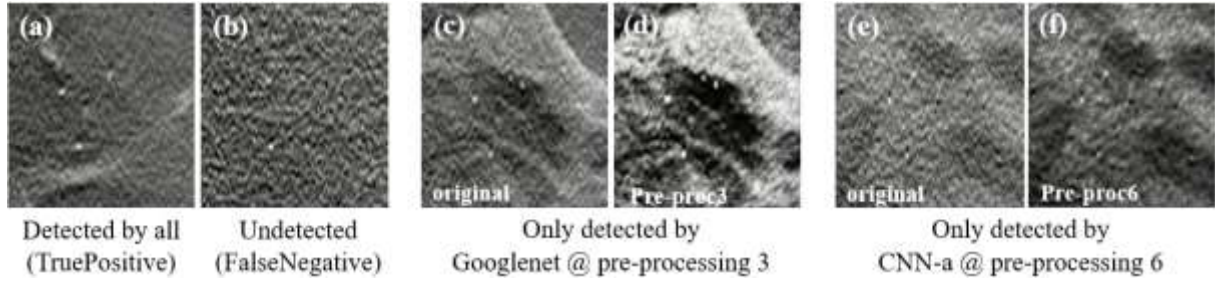


Figure 6.9. Some examples of MCs in the DBT data used. (a) True positive (case classified correctly as positive by all CNNs, even in the original image); (b) False negative (case incorrectly classified as negative by all CNNs, even varying the pre-processing); (c) original case classified as negative and that was only detected by Googlenet when pre-processed with method 3 (d); (e) original case classified as negative and that was only detected by CNN-a when pre-processed with method 6 (f).

The variations and differences in AUC values obtained for each situation were, in general, in agreement with the specificity, sensitivity and accuracy values obtained in Figure 6.7. Although specificity values are higher than sensitivity in most cases, these differences are not significant ($p\text{-value} > 0.05$ in all cases). As for accuracy, Googlenet and CNN-a presented the best values of 85.68% and 82.45%, respectively.

In the VICTRE database, it is possible to separate the cases by breast density and a study was published where a model observer was trained separately for detecting lesions in each of the four breast density types and then tested it on the same density type to obtain the individual AUC for each density [88]. As a conclusion of this study, Zeng et al. believe it would be appropriate to train the model observer with mixed breast density images. This was exactly what we did with the deep learning architectures proposed in this work. However, in order to understand whether the presented methodologies were influenced or not by breast density, the same CNNs were tested separately for classifying the DBT data about the presence of MCs in each of the four breast density types (fatty, scattered, heterogeneous and dense) and the results analyzed in terms of AUC. As seen in Figure 6.7, only Squeezenet was especially sensitive to density, showing significant differences in detection between three density types. The correct classification of cases with MCs in dense breasts with the Squeezenet was significantly lower compared to the other densities. In general, due to the lower anatomical background, fatty

breasts allowed good classifications of cases with MCs. The Googlenet was the exception, with fatty breasts corresponding to the lowest AUC value ($p>0.05$).

Training and inference times of Table 6.5 are purely indicative as they vary depending on the computation power available. But, in relative terms, the already existent networks (Googlenet, Resnet18, Squeezenet and Alexnet) lead with the four longest times. On the other hand, although the CNN with the best AUC (Googlenet) showed the longest time, the second best (CNN-a) was the faster network. As inference time is the key when the models are used in clinic, it should be noted that with CNN-a it was possible to classify an image never seen by the model before about three times faster than with Googlenet. From our point of view, this fact makes this architecture adapted from Alexnet very interesting for future studies that involve more complex and longer trainings, such as object detection with state-of-the-art faster region-based CNNs. One of the most determining factors in the training/testing time of these CNNs is the feature extraction network that is used as the basis. Thus, a faster model such as CNN-a, which presents good results regarding the classification of cases with MCs, should be an option to be studied in the future.

In two published studies (2D and 3D), where a prescreening stage generates possible MCs and the proposed CNNs differentiate between true MCs and false positives, AUCs values of 93% [51] and 97.65% [69] were reported. Both studies used ROIs instead of the whole image/volume. Some regions don't have any lesion or relevant information, while others contain only the lesions. On the other hand, in a study where the main objective was to compare the detection of MCs in images reconstructed with two different reconstruction algorithms (EMPIRE and filtered back projection), small 3D patches were used as input and the best result obtained in terms of partial AUC was 88.0% [59].

In another study, a ROI was selected for each lesion on a DBT key slice, features were extracted using a pre-trained CNN and served as input to a support vector machine classifier trained in the task of predicting likelihood of malignancy [63]. The AUC results obtained in CC view for MCs detection was 82%. Other views were included and, considering MLO (mediolateral oblique) in addition to CC view, AUC improved to 97%, showing the importance of having both views available.

The only work that took the whole image information into account, used 2D synthetic mammographic images obtained from DBT exams to train a multi-view deep CNN to classify screening images into BI RADS classes (0: further evaluation is required due to a suspicious

abnormality; 1: the mammogram is negative; and 2: the mammogram is benign). The AUC values obtained were as follow: BI-RADS 0 vs. others: 91.2%; BI-RADS 1 vs. others: 90.5%; BIRADS 2 vs. others: 90.0% [67].

A direct comparison between literature values and those obtained in this work is not fair due to several reasons. The first is that different databases were used (those of the studies mentioned are all private databases). The second is that the training data have quite different characteristics due to different detection tasks. Some used only small parts of the data, and those which used the entire image did not refer to DBT slices but rather to synthetic mammograms obtained with DBT. Nevertheless, it is possible to confirm that the results obtained by our study (maximum value of AUC achieved: 94.19%) are quite competitive when compared to those available in the literature.

There are some limitations in this study. The first is that the available dataset is limited to the CC view and one manufacturer. The second is that only one type of lesion (MCs) was considered and, within the available data, there may be some similarities between lesions. We tried to overcome this fact through data augmentation with reflection and rotation. The third is that, despite being very realistic, the data are simulated and therefore don't correspond to real patients. Finally, since DBT is a 3D technique, the fact that we consider information in 2D slices can limit the advantage provided by the depth information. Furthermore, the true clinical value lies on the classification of a volume, because this is what radiologists do every day in clinical practice. We believe that this work is a starting point and can serve as a basis for the implementation of a 3D training with all volume and 3D architectures, considering real data volumes and not just some slices. In addition, it will also be important to diversify the lesions, including data obtained from other views (MLO), manufacturers and reconstruction algorithms. As for the training of the CNNs themselves, other optimizers that have been producing good results (such as Adam optimizer), different mini-batch sizes and learning rates should be tested and evaluated.

6.5. Conclusions

Deep learning AI algorithms composed of multi-layered CNNs have been growing over the past five years and have shown very promising results in supporting the detection of breast cancer. One of the great difficulties in training these algorithms is the lack of labelled DBT

databases. Furthermore, all published studies refer to private databases, thus limiting the comparison and improvement of the works carried out.

In this study, a public DBT dataset was used to train from scratch four popular CNNs and a new CNN model proposed by us. The main task of our algorithms was to classify a DBT case for the presence or absence of MCs, given the whole DBT image as input. Beside the original data, six different pre-processing methodologies, which the main purpose was to highlight MCs, were implemented to generate different input datasets.

Classifying the whole image according to the presence or absence of MCs is a difficult task due to the size of MCs and all the information present in an entire image. With this work, we were able to achieve encouraging outcomes in this regard, obtaining similar results to other studies for the detection of larger lesions such as masses. The classification of cases with/without MCs was greatly influenced by the type of input data and our new model achieved the second best performance in the shortest time, thus becoming a very interesting model to be considered in future studies.

ACKNOWLEDGMENTS

This work was supported by Universidade de Lisboa (PhD grant) and Fundação para a Ciência e Tecnologia – Portugal (Grant No. SFRH/BD/135733/2018 and FCT-IBEB Strategic Project UIDB/00645/2020).

References

1. Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries*. CA Cancer J Clin, 2021. **71**(3): p. 209-249.
2. Siegel, R.L., et al., *Cancer Statistics, 2021*. CA Cancer J Clin, 2021. **71**(1): p. 7-33.
3. Tabár, L., et al., *The incidence of fatal breast cancer measures the increased effectiveness of therapy in women participating in mammography screening*. Cancer, 2019. **125**(4): p. 515-523.
4. Skaane, P., et al., *Performance of breast cancer screening using digital breast tomosynthesis: results from the prospective population-based Oslo Tomosynthesis Screening Trial*. Breast Cancer Research and Treatment, 2018. **169**(3): p. 489-496.
5. Ciatto, S., et al., *Integration of 3D digital mammography with tomosynthesis for population breast-cancer screening (STORM): a prospective comparison study*. The Lancet Oncology, 2013. **14**(7): p. 583-589.
6. Haas, B.M., et al., *Comparison of Tomosynthesis Plus Digital Mammography and Digital Mammography Alone for Breast Cancer Screening*. Radiology, 2013. **269**(3): p. 694-700.
7. Rose, S.L., et al., *Implementation of Breast Tomosynthesis in a Routine Screening Practice: An Observational Study*. American Journal of Roentgenology, 2013. **200**(6): p. 1401-1408.
8. Greenberg, J.S., et al., *Clinical Performance Metrics of 3D Digital Breast Tomosynthesis Compared With 2D Digital Mammography for Breast Cancer Screening in Community Practice*. American Journal of Roentgenology, 2014. **203**(3): p. 687-693.

9. McDonald, E.S., et al., *Effectiveness of Digital Breast Tomosynthesis Compared With Digital Mammography: Outcomes Analysis From 3 Years of Breast Cancer Screening*. JAMA Oncology, 2016. **2**(6): p. 737-743.
10. Zackrisson, S., et al., *One-view breast tomosynthesis versus two-view mammography in the Malmö Breast Tomosynthesis Screening Trial (MBTST): a prospective, population-based, diagnostic accuracy study*. The Lancet Oncology, 2018. **19**(11): p. 1493-1503.
11. Bernardi, D., et al., *Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study*. The Lancet Oncology, 2016. **17**(8): p. 1105-1113.
12. Lång, K., et al., *Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmö Breast Tomosynthesis Screening Trial, a population-based study*. Eur Radiol, 2016. **26**(1): p. 184-190.
13. Gilbert, F.J., et al., *Accuracy of Digital Breast Tomosynthesis for Depicting Breast Cancer Subgroups in a UK Retrospective Reading Study (TOMMY Trial)*. Radiology, 2015. **277**(3): p. 697-706.
14. Hofvind, S., et al., *Digital Breast Tomosynthesis and Synthetic 2D Mammography versus Digital Mammography: Evaluation in a Population-based Screening Program*. Radiology, 2018. **287**(3): p. 787-794.
15. Freer, P.E., et al., *Clinical implementation of synthesized mammography with digital breast tomosynthesis in a routine clinical practice*. Breast Cancer Research and Treatment, 2017. **166**(2): p. 501-509.
16. Food and Drug Administration (FDA) U.S. . *Premarket Approval application supplement for the Selenia Dimensions 3D System*. 2013 [cited 2021 May].
17. Skaane, P., et al., *Comparison of Digital Mammography Alone and Digital Mammography Plus Tomosynthesis in a Population-based Screening Program*. Radiology, 2013. **267**(1): p. 47-56.
18. Tagliafico, A.S., et al., *Accuracy and reading time for six strategies using digital breast tomosynthesis in women with mammographically negative dense breasts*. Eur Radiol, 2017. **27**(12): p. 5179-5184.
19. Balleyguier, C., et al., *Improving digital breast tomosynthesis reading time: A pilot multi-reader, multi-case study using concurrent Computer-Aided Detection (CAD)*. Eur J Radiol, 2017. **97**: p. 83-89.
20. Benedikt, R.A., et al., *Concurrent Computer-Aided Detection Improves Reading Time of Digital Breast Tomosynthesis and Maintains Interpretation Performance in a Multireader Multicase Study*. American Journal of Roentgenology, 2017. **210**(3): p. 685-694.
21. Chae, E.Y., et al., *Decrease in interpretation time for both novice and experienced readers using a concurrent computer-aided detection system for digital breast tomosynthesis*. Eur Radiol, 2019. **29**(5): p. 2518-2525.
22. Poplack, S.P., et al., *Digital breast tomosynthesis: initial experience in 98 women with abnormal digital screening mammography*. AJR Am J Roentgenol, 2007. **189**(3): p. 616-23.
23. Andersson, I., et al., *Breast tomosynthesis and digital mammography: a comparison of breast cancer visibility and BIRADS classification in a population of cancers with subtle mammographic findings*. Eur Radiol, 2008. **18**(12): p. 2817-25.
24. Spangler, M.L., et al., *Detection and Classification of Calcifications on Digital Breast Tomosynthesis and 2D Digital Mammography: A Comparison*. American Journal of Roentgenology, 2011. **196**(2): p. 320-324.
25. Kopans, D., et al., *Calcifications in the breast and digital breast tomosynthesis*. Breast J, 2011. **17**(6): p. 638-44.
26. Svane, G., et al., *Clinical experience of photon counting breast tomosynthesis: comparison with traditional mammography*. Acta Radiologica, 2011. **52**(2): p. 134-142.
27. Wallis, M.G., et al., *Two-View and Single-View Tomosynthesis versus Full-Field Digital Mammography: High-Resolution X-Ray Imaging Observer Study*. Radiology, 2012. **262**(3): p. 788-796.

28. Nyante, S.J., et al., *The association between mammographic calcifications and breast cancer prognostic factors in a population-based registry cohort*. Cancer, 2017. **123**(2): p. 219-227.
29. D'Orsi, C.J., *2013 ACR BI-RADS Atlas: Breast Imaging Reporting and Data System* 2014: American College of Radiology.
30. Samala, R.K., et al., *Digital breast tomosynthesis: Computer-aided detection of clustered microcalcifications on planar projection images*. Phys Med Biol, 2014. **59**(23): p. 7457-7477.
31. Samala, R.K., et al., *Analysis of computer-aided detection techniques and signal characteristics for clustered microcalcifications on digital mammography and digital breast tomosynthesis*. Phys Med Biol, 2016. **61**(19): p. 7092-7112.
32. Park, S.C., et al., *Applying a 2D based CAD scheme for detecting micro-calcification clusters using digital breast tomosynthesis images: an assessment*. 2008: p. 691507-691507.
33. Reiser, I., et al., *Automated detection of microcalcification clusters for digital breast tomosynthesis using projection data only: A preliminary study*. Med Phys, 2008. **35**(4): p. 1486-1493.
34. Bernard, S., S. Muller, and J. Onativia, *Computer-Aided Microcalcification Detection on Digital Breast Tomosynthesis Data: A Preliminary Evaluation*, in *Digital Mammography: 9th International Workshop, IWDM 2008 Tucson, AZ, USA, July 20-23, 2008 Proceedings*, E.A. Krupinski, Editor 2008, Springer Berlin Heidelberg: Berlin, Heidelberg. p. 151-157.
35. Sahiner, B., et al., *Computer-aided detection of clustered microcalcifications in digital breast tomosynthesis: a 3D approach*. Med Phys, 2012. **39**(1): p. 28-39.
36. Samala, R.K., et al., *Computer-aided detection of clustered microcalcifications in multiscale bilateral filtering regularized reconstructed digital breast tomosynthesis volume*. Med Phys, 2014. **41**(2): p. 021901-n/a.
37. Wei, J., et al., *Multichannel response analysis on 2D projection views for detection of clustered microcalcifications in digital breast tomosynthesis*. Med Phys, 2014. **41**(4): p. 041913-041913.
38. Samala, R.K., et al., *Computer-aided detection system for clustered microcalcifications in digital breast tomosynthesis using joint information from volumetric and planar projection images*. Phys Med Biol, 2015. **60**(21): p. 8457-8479.
39. Fenton, J.J., et al., *Influence of Computer-Aided Detection on Performance of Screening Mammography*. New England Journal of Medicine, 2007. **356**(14): p. 1399-1409.
40. Lehman, C.D., et al., *Diagnostic Accuracy of Digital Screening Mammography With and Without Computer-Aided Detection*. JAMA Internal Medicine, 2015. **175**(11): p. 1828-1837.
41. Katzen, J. and K. Dodelzon, *A review of computer aided detection in mammography*. Clinical Imaging, 2018. **52**: p. 305-309.
42. Sechopoulos, I., J. Teuwen, and R. Mann, *Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art*. Seminars in Cancer Biology, 2020.
43. Rodriguez-Ruiz, A., et al., *Stand-Alone Artificial Intelligence for Breast Cancer Detection in Mammography: Comparison With 101 Radiologists*. Journal of the National Cancer Institute, 2019. **111**(9): p. 916-922.
44. McKinney, S.M., et al., *International evaluation of an AI system for breast cancer screening*. Nature, 2020. **577**(7788): p. 89-94.
45. Kim, H.-E., et al., *Changes in cancer detection and false-positive recall in mammography using artificial intelligence: a retrospective, multireader study*. The Lancet Digital Health, 2020. **2**(3): p. e138-e148.
46. Wang, X., et al., *Inconsistent Performance of Deep Learning Models on Mammogram Classification*. Journal of the American College of Radiology, 2020. **17**(6): p. 796-803.
47. Schaffter, T., et al., *Evaluation of Combined Artificial Intelligence and Radiologist Assessment to Interpret Screening Mammograms*. JAMA Network Open, 2020. **3**(3): p. e200265-e200265.
48. Rodríguez-Ruiz, A., et al., *Detection of Breast Cancer with Mammography: Effect of an Artificial Intelligence Support System*. Radiology, 2019. **290**(2): p. 305-314.

49. Conant, E.F., et al., *Improving Accuracy and Efficiency with Concurrent Use of Artificial Intelligence for Digital Breast Tomosynthesis*. Radiology: Artificial Intelligence, 2019. **1**(4): p. e180096.
50. van Winkel, S.L., et al., *Impact of artificial intelligence support on accuracy and reading time in breast tomosynthesis image interpretation: a multi-reader multi-case study*. Eur Radiol, 2021.
51. Samala, R., et al., *Deep-learning convolution neural network for computer-aided detection of microcalcifications in digital breast tomosynthesis*. SPIE Medical Imaging. Vol. 9785. 2016: SPIE.
52. Fotin, S., et al., *Detection of soft tissue densities from digital breast tomosynthesis: comparison of conventional and deep learning approaches*. SPIE Medical Imaging. Vol. 9785. 2016: SPIE.
53. Kim, D.H., S.T. Kim, and Y.M. Ro. *Latent feature representation with 3-D multi-view deep convolutional neural network for bilateral analysis in digital breast tomosynthesis*. in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. 2016.
54. Samala, R.K., et al., *Mass detection in digital breast tomosynthesis: Deep convolutional neural network with transfer learning from mammography*. Med Phys, 2016. **43**(12): p. 6654-6654.
55. Kim, D.H., et al., *Latent feature representation with depth directional long-term recurrent learning for breast masses in digital breast tomosynthesis*. Phys Med Biol, 2017. **62**(3): p. 1009-1031.
56. Zhang, X., et al., *Classification of Whole Mammogram and Tomosynthesis Images Using Deep Convolutional Neural Networks*. IEEE Transactions on NanoBioscience, 2018. **17**(3): p. 237-242.
57. Samala, R.K., et al., *Evolutionary pruning of transfer learned deep convolutional neural network for breast cancer diagnosis in digital breast tomosynthesis*. Physics in Medicine & Biology, 2018. **63**(9): p. 095005.
58. Yousefi, M., A. Krzyżak, and C.Y. Suen, *Mass detection in digital breast tomosynthesis data using convolutional neural networks and multiple instance learning*. Computers in Biology and Medicine, 2018. **96**: p. 283-293.
59. Rodriguez-Ruiz, A., et al., *New reconstruction algorithm for digital breast tomosynthesis: better image quality for humans and computers*. Acta radiologica (Stockholm, Sweden : 1987), 2018. **59**(9): p. 1051-1059.
60. Mordang, J.-J., et al. *Automatic Microcalcification Detection in Multi-vendor Mammography Using Convolutional Neural Networks*. 2016. Cham: Springer International Publishing.
61. Zhang, Y., et al. *2D Convolutional Neural Networks for 3D Digital Breast Tomosynthesis Classification*. in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. 2019.
62. Liang, G., et al. *Joint 2D-3D Breast Cancer Classification*. in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. 2019.
63. Mendel, K., et al., *Transfer Learning From Convolutional Neural Networks for Computer-Aided Diagnosis: A Comparison of Digital Breast Tomosynthesis and Full-Field Digital Mammography*. Acad Radiol, 2019. **26**(6): p. 735-743.
64. Singh, S., et al. *Adaptation of a deep learning malignancy model from full-field digital mammography to digital breast tomosynthesis*. 2020. arXiv:2001.08381.
65. Li, X., et al., *Digital breast tomosynthesis versus digital mammography: integration of image modalities enhances deep learning-based breast mass classification*. Eur Radiol, 2020. **30**(2): p. 778-788.
66. Wang, L., et al., *Multi-path synergic fusion deep neural network framework for breast mass classification using digital breast tomosynthesis*. Physics in Medicine & Biology, 2020. **65**(23): p. 235045.
67. Seyyedi, S., et al., *SCREENet: A Multi-view Deep Convolutional Neural Network for Classification of High-resolution Synthetic Mammographic Screening Scans*. ArXiv, 2020. **abs/2009.08563**.

68. Matthews, T.P., et al., *A Multisite Study of a Breast Density Deep Learning Model for Full-Field Digital Mammography and Synthetic Mammography*. Radiology: Artificial Intelligence, 2021. **3**(1): p. e200015.
69. Zheng, J., et al., *3D Context-Aware Convolutional Neural Network for False Positive Reduction in Clustered Microcalcifications Detection*. IEEE Journal of Biomedical and Health Informatics, 2021. **25**(3): p. 764-773.
70. R V, A., A. R., and S. A P, *Augmenting Transfer Learning with Feature Extraction Techniques for Limited Breast Imaging Datasets*. J Digit Imaging, 2021.
71. Bai, J., et al., *Applying deep learning in digital breast tomosynthesis for automatic breast cancer detection: A review*. Med Image Anal, 2021. **71**: p. 102049.
72. Buda, M., et al. *Detection of masses and architectural distortions in digital breast tomosynthesis: a publicly available dataset of 5,060 patients and a deep learning model*. 2020. arXiv:2011.07995.
73. Badano, A., et al., *Evaluation of Digital Breast Tomosynthesis as Replacement of Full-Field Digital Mammography Using an In Silico Imaging Trial*. JAMA Network Open, 2018. **1**(7): p. e185474-e185474.
74. VICTRE. *The VICTRE Trial: Open-Source, In-Silico Clinical Trial For Evaluating Digital Breast Tomosynthesis*. 2018 [cited 2021; Available from: <https://wiki.cancerimagingarchive.net/display/Public/The+VICTRE+Trial%3A+Open-Source%2C+In-Silico+Clinical+Trial+For+Evaluating+Digital+Breast+Tomosynthesis>].
75. Sidky, E.Y., et al., *Enhanced imaging of microcalcifications in digital breast tomosynthesis through improved image-reconstruction algorithms*. Med Phys, 2009. **36**(11): p. 4920-32.
76. Lu, Y., et al., *Selective-diffusion regularization for enhancement of microcalcifications in digital breast tomosynthesis reconstruction*. Med Phys, 2010. **37**(11): p. 6003-6014.
77. Mota, A.M., et al., *Total variation minimization filter for DBT imaging*. Med Phys, 2015. **42**(6Part1): p. 2827-2836.
78. Michielsen, K., et al., *Design of a model observer to evaluate calcification detectability in breast tomosynthesis and application to smoothing prior optimization*. Med Phys, 2016. **43**(12): p. 6577-6587.
79. Mota, A.M., et al., *An Enhanced Visualization of DBT Imaging Using Blind Deconvolution and Total Variation Minimization Regularization*. IEEE Transactions on Medical Imaging, 2020. **39**(12): p. 4094-4101.
80. Zuiderveld, K., *Contrast limited adaptive histogram equalization*, in *Graphics gems IV* 1994, Academic Press Professional, Inc. p. 474-485.
81. MathWorks. *MATLAB adapthisteq function*. 2021 [cited 2021 May]; Available from: <https://www.mathworks.com/help/images/ref/adapthisteq.html>.
82. Krizhevsky, A., I. Sutskever, and G.E. Hinton, *Imagenet classification with deep convolutional neural networks*. Advances in neural information processing systems, 2012. **25**: p. 1097-1105.
83. Szegedy, C., et al. *Going deeper with convolutions*. in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2015.
84. He, K., et al. *Deep Residual Learning for Image Recognition*. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016.
85. Iandola, F.N., et al., *SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size*. arXiv:1602.07360, 2016.
86. MathWorks. *Transfer Learning*. 2021 [cited 2021 May]; Available from: <https://www.mathworks.com/discovery/transfer-learning.html>.
87. Vourtsis, A. and W.A. Berg, *Breast density implications and supplemental screening*. Eur Radiol, 2019. **29**(4): p. 1762-1777.
88. Zeng, R., et al., *Computational reader design and statistical performance evaluation of an in-silico imaging clinical trial comparing digital breast tomosynthesis with full-field digital mammography*. Journal of medical imaging (Bellingham, Wash.), 2020. **7**(4): p. 042802-042802.

Detection of microcalcifications in Digital Breast Tomosynthesis using Faster R-CNN and 3D volume rendering

7

Ana M. Mota^{a*}, Matthew J. Clarkson^b, Pedro Almeida^a and Nuno Matela^a, “Detection of microcalcifications in Digital Breast Tomosynthesis using Faster R-CNN and 3D volume rendering”. Accepted as an extended conference paper in *Bioimaging 2022*: 9th International Conference on Bioimaging.

BioImaging is part of BIOSTEC, the International Joint Conference on Biomedical Engineering Systems and Technologies.

^aFaculdade de Ciências, Instituto de Biofísica e Engenharia Biomédica, Universidade de Lisboa, Campo Grande 1749-016 Lisboa, Portugal;

^bDepartment of Medical Physics and Biomedical Engineering and the Centre for Medical Image Computing, University College London, London, UK.

* Corresponding author: ammota@fc.ul.pt

Abstract

Microcalcification clusters (MCs) are one of the most important biomarkers for breast cancer and Digital Breast Tomosynthesis (DBT) has consolidated its role in breast cancer imaging. As there are mixed observations about MCs detection using DBT, it is important to develop tools that improve this task. Furthermore, the visualization mode of MCs is also crucial, as their diagnosis is associated with their 3D morphology. In this work, DBT data from a public database were used to train a faster region-based convolutional neural network (R-CNN) to locate MCs in entire DBT. Additionally, the detected MCs were further analyzed through standard 2D visualization and 3D volume rendering (VR) specifically developed for DBT data. For MCs detection, the sensitivity of our Faster R-CNN was 60% with 4 false positives. These preliminary results are very promising and can be further improved. On the other hand, the 3D VR visualization provided important information, with higher quality and discernment of the detected MCs. The developed pipeline may help radiologists since (1) it indicates specific breast regions with possible lesions that deserve additional attention and (2) as the rendering of the MCs is similar to a segmentation, a detailed complementary analysis of their 3D morphology is possible.

6.1. Introduction

Breast cancer is the type of cancer with higher incidence, among all cancers and both sexes, and it still represents the biggest cause of cancer mortality among women [1]. The mortality rate from this disease has been decreasing in the last decades due to the new therapies and the implementation of screening programs for early detection [2].

The use of Digital Breast Tomosynthesis (DBT) has been confirming its potential to address the tissue overlapping limitations of Digital Mammography (DM), the gold standard for breast screening until recently. In fact, by including synthetic mammographies generated from DBT data, DBT alone is now used as a stand-alone modality to replace DM [3-9]. DBT volume data can be analyzed in depth through several 2D slices (standard visualization slice-by-slice). This multi-slice inspection leads to a longer analysis time (because instead of two images, radiologists have to inspect an average of sixty images per patient), which represents a problem in daily practice and screening environment [10-12].

Computer-Aided Detection (CAD) systems based on DBT have been implemented and evaluated in an attempt to shorten the reading time while maintaining the radiologist performance. However, despite the efforts and improvements already achieved, due to the high false positive (FP) rates and low specificity, these CAD systems have not reached a level of performance that can be translated into a true improvement in the real screening of breast cancer [13-16].

On the other hand, a different type of visualization, such as 3D volume rendering (VR), may play an important complementary role in breast cancer diagnosis [17]. With a visualization of the object through multiple angles, one of the advantages of VR is to provide an intuitive understanding of the underlying data at once. In addition, as VR yields a true depth perception [18], it can help in the analysis of lesions such as microcalcification clusters (MCs), sometimes referred as harder to detect in DBT. These MCs are often spread across several slices in the slice-by-slice visualization, making the interpretation difficult. In this way, a better understanding of its true 3D morphology is important to differentiate between benign and malignant microcalcifications.

In recent years, the increase in computational power and bigger datasets have allowed the development of algorithms for automatic object detection with deep learning. The region-based convolutional neural networks (R-CNNs) are one of the main current focuses of research and

development of these methods [19]. As R-CNN and its descendent “fast R-CNN” [20] are computationally expensive and extremely slow, another method has emerged: “Faster R-CNN” [21]. With this object detection network, both the CNN-based regional proposals and the regional classification module are trained together with significant weight sharing, led to increased sensitivity for object detection and faster speed.

The published studies that use deep CNNs to detect and localize lesions in DBT are still very limited. In fact, the few works that exist are related with the detection of soft tissue lesions [22-25]. Regarding the use of Faster R-CNN in particular, [26] developed a CAD system for masses detection in DBT using a Faster R-CNN, which is later compared to a framework of a 3D-Mask R-CNN for mass detection and segmentation [27]. [28] propose a Faster R-CNN that uses mammary gland distribution as a prior information to detect architectural distortions in DBT.

In this paper, a Faster R-CNN was trained for detecting MCs in DBT. The aim is to input a whole DBT image into the network and have a direct answer about the localization or absence of MCs. This information about the location is then introduced into a 3D VR visualization software so that a 3D volume of interest containing the predicted MCs can be obtained. A public simulated database was used and the preliminary results obtained are presented. To the best of our knowledge, this is the first study of automatic localization of MCs in whole DBT images and the first time the DBT output of a deep CNN is rendered and presented as a 3D volume of interest.

7.2. Materials and methods

This work was implemented on the MATLAB R2020a and a NVIDIA Quadro P4000 GPU computer was used.

7.2.1. Database and pre-processing

The public database of Virtual Imaging Clinical Trial for Regulatory Evaluation (VICTRE) project which contains a total of 2986 virtual realistic patients imaged with DBT was used [29, 30]. This database contains cases without lesions (absent) and with malignant masses and MCs. For training, only cases with MCs were considered (915 in total: 665 complete breast images and 250 images containing only MCs) and for the testing, absent and MC cases were included (280 and 284 complete breast images, respectively). Each case with lesion contains four MCs consisting of 5 calcified lesions modelled as 195, 179, and 171 μm of solid calcium oxalate.

In addition to the information about the presence or absence of MCs, in cases where MCs were present, information about the corresponding bounding boxes (BBs) was also given to the network. This information, in the form of x, y and z coordinates as well as width and height, is in the VICTRE database.

We adopted the usual distribution of breast density in the general population: 10% fatty, 40% scattered, 40% heterogeneous and 10% dense. The reconstructed cases have different dimension in x, y and z, depending on breast density: $1624 \times 1324 \times 62$, $1421 \times 1024 \times 57$, $1148 \times 753 \times 47$ and $1130 \times 477 \times 38$ for fatty, scattered, heterogeneous and dense breasts, respectively, with a voxel size of $0.085 \times 0.085 \times 1 \text{ mm}^3$.

The data intensity was first normalized between 0 and 1 and then squared to highlight the higher intensity values belonging to the MCs, while attenuate the lower ones. With this pre-processing step our aim was to specifically increase the contrast of regions of higher intensities. In addition, through binarization and region growing operations, binary masks that keep information belonging to the breast and make everything else zero were created (background suppression).

7.2.2. Faster R-CNN object detector

Faster R-CNN is based on a CNN and a region proposal network (RPN) for detecting, localizing and classifying objects in an image. The CNN module (typically a pre-trained CNN), outputs a set of feature maps and, for that reason, it is also called feature extraction network. In our work, we used the ResNet-18 model, trained on more than a million images from the ImageNet database [31]. The RPN is on top of the last convolutional layer of the CNN and it uses default bounding boxes (anchors) with different sizes and aspect ratios over the feature maps generated from pre-trained CNN in order to find objects with varying sizes and shapes. It is trained to output a set of object proposals on the image, each with an “objectness” score, regardless of the class of the object (it only looks if it is an object or background). The boxes with the highest score are called region proposals and are introduced in another branch of the network where they are resampled to a fixed size (ROI Pooling) and, typically using few fully connected layers, the class of the object present in the boundary boxes is determined. Further details about Faster R-CNN can be found in the original paper [21]. The main parameters used to define our Faster R-CNN are presented in Table 7.1.

Table 7.1. Parameters used to design the Faster R-CNN.

Input size	224x224x3
Anchor Boxes	42x27; 63x45; 45x41
Pre-trained CNN	ResNet-18

7.2.3. Faster R-CNN training

The Faster R-CNN was trained using the end-to-end method, where the RPN and the region classification networks were trained simultaneously along 660k iterations. Table 7.2 presents the main training options defined for this work.

During training, several regions of the image are processed from the training database. The positive and negative overlap range properties control which image regions are used for training. This overlap ratio is defined as the Intersection over Union (IoU) metric that describes the extent of overlap between two boxes (ground truth and predicted BB). The greater the region of overlap, the greater the IOU. The model was trained to minimize the mean square error loss between the predicted BBs and the ground truth using the Stochastic Gradient Descent optimizer [32].

Table 7.2. Options used to train the Faster R-CNN.

Solver	Stochastic Gradient Descent with momentum
Momentum	0.9
Size of mini-batch	1
Learning rate	1e-3
Factor for L_2 regularization	5e-4
Training method	End-to-end
Positive Overlap Range	[0.3 1]
Negative Overlap Range	[0 0.1]

To prevent overfitting, each image in the training set was augmented by random reflection in the left-right direction and rotation between -20° and 20° . In addition, a L_2 regularization term for the weight decay was introduced in the loss function.

7.2.4. Evaluation metrics

The network's ability to accurately detect and locate the MCs was evaluated through the Free-response Receiver Operating Characteristic (FROC) curve [33]. To obtain a point on the FROC curve, a threshold value is fixed and only the findings that have scores above that threshold are selected. Then the sensitivity (true positive fraction) and mean number of FPs per image are determined.

7.2.5. Data visualization

Figure 7.1 shows the scheme followed during and after Faster R-CNN training. A testing set is evaluated for the detection of MCs using the trained Faster R-CNN and the output results (predicted BBs) are visualized. In addition to the standard 2D visualization, the output detection was also analyzed through 3D visualization with VR. The 2D visualization was performed by calculating the 2D maximum intensity projection (MIP) considering the slice where the cluster was detected and the four adjacent slices (two down and two up). The 3D visualization was performed through VR with 3D MIP considering the same slices.

The Visualization Toolkit library (VTK) version 7.1.0. (Kitware, New York, EUA) [34, 35] was used to develop 3D specific software in order to visualize DBT data through VR. The opacity/color transfer functions for an adequate rendering of these data were calculated accordingly to previous work [36].

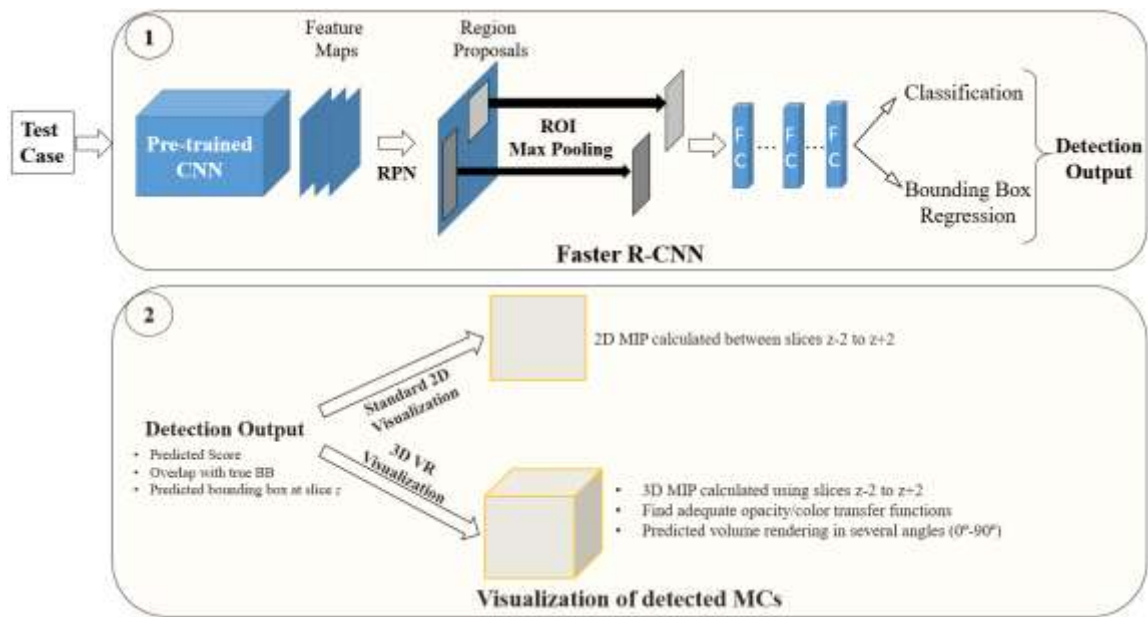


Figure 7.1. Pipeline followed for connection between the output of the trained faster R-CNN and the volume rendering visualization of the detected object.

In VR, changing the azimuth of a camera rotates its position around the focal point [35] allowing an immediate notion of the entire volume in 3D. In this way, the volume of interest containing the detected MCs is presented from several angles (from 0° to 90°).

7.3. Results

The training of 660k iterations was performed during 12 days. The analysis of one test image was done in 0.6 seconds (mean time) and for an entire DBT volume our faster R-CNN needed, on average, 29 seconds (depending on the size).

7.3.1. Faster R-CNN detection

Figure 7.2 presents the FROC curve for the performance of the training model to accurately detect and locate the MCs for several thresholds. In addition, the discriminative sensitivity values obtained for less than 8 FP /image are detailed in the Table 7.3.

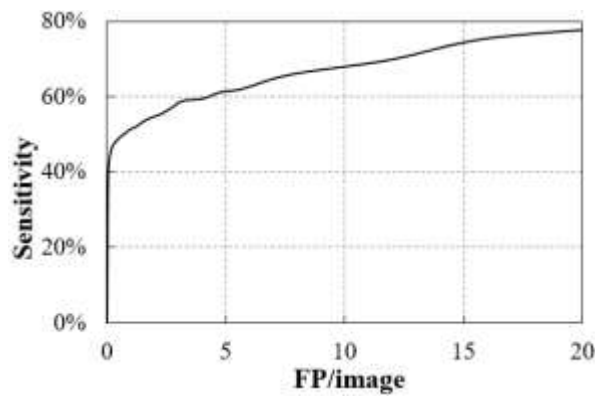


Figure 7.2. The FROC curve for the test dataset.

Table 7.3. The sensitivity values for less than 8 FP/image.

Sensitivity (%)	FP/image	# MCs detected	# MCs undetected
40	0.1	125	159
47	0.2	146	138
51	0.8	158	126
54	1.8	170	114
57	2.7	178	106
59	3.2	184	100
61	4.8	186	98
62	5.7	194	90
66	7.8	206	78

7.3.2. Data visualization

Four examples of detection output, including the FPs (yellow) and true positives (green) BBs, obtained with a threshold of 0.9 are presented in Figure 7.3. The corresponding score is also shown. As described, each detected MC is presented through two visualization modes: 2D slice-by-slice and 3D VR. As 3D VR is inspected through several angles (0, 22.5°, 45°, 67.5° and

90°), 2D MIP slice-by-slice is presented using xy and xz representations for comparison with VR 0° and 90°, respectively.

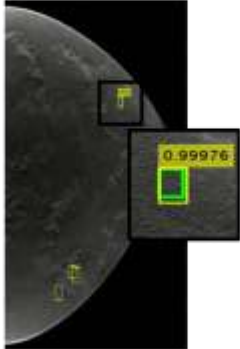


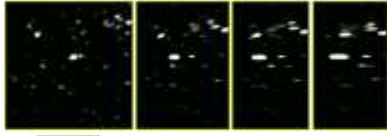

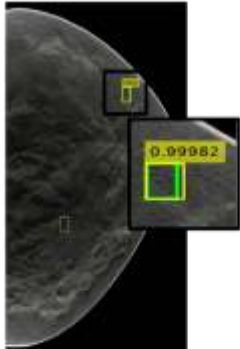
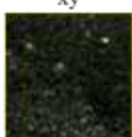

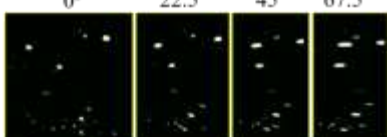

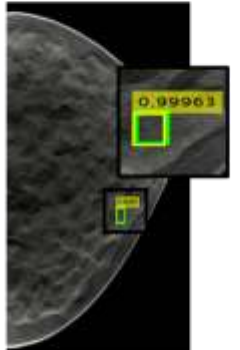
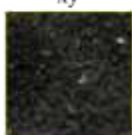

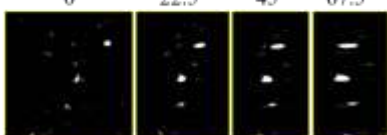

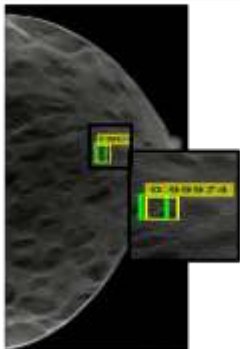
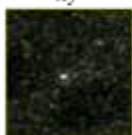

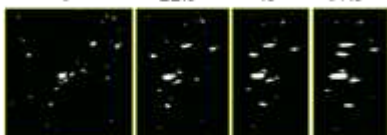

Faster R-CNN detection (threshold: 0.90)	Visualization mode	
	2D Slice-by-slice	3D Volume rendering
	xy  xz 	0° 22.5° 45° 67.5°  90° 
	xy  xz 	0° 22.5° 45° 67.5°  90° 
	xy  xz 	0° 22.5° 45° 67.5°  90° 
	xy  xz 	0° 22.5° 45° 67.5°  90° 

Figure 7.3. Example of four detection outputs obtained with a threshold of 0.9. Green: Ground truth BB; Yellow: predicted BB (without score: FPs, with score: true positives). The predicted results are visualized with 2D slice-by-slice represented through xy and xz planes and 3D VR with five different angles (0°, 22.5°, 45°, 67.5° and 90°).

In Figure 7.4 (a) are presented four examples of missed detections (false negatives) and on Figure 7.4 (b) four incorrect detections (FPs). The detection results are then visualized through 2D slice-by-slice and 3D VR at xy and 0° , respectively.

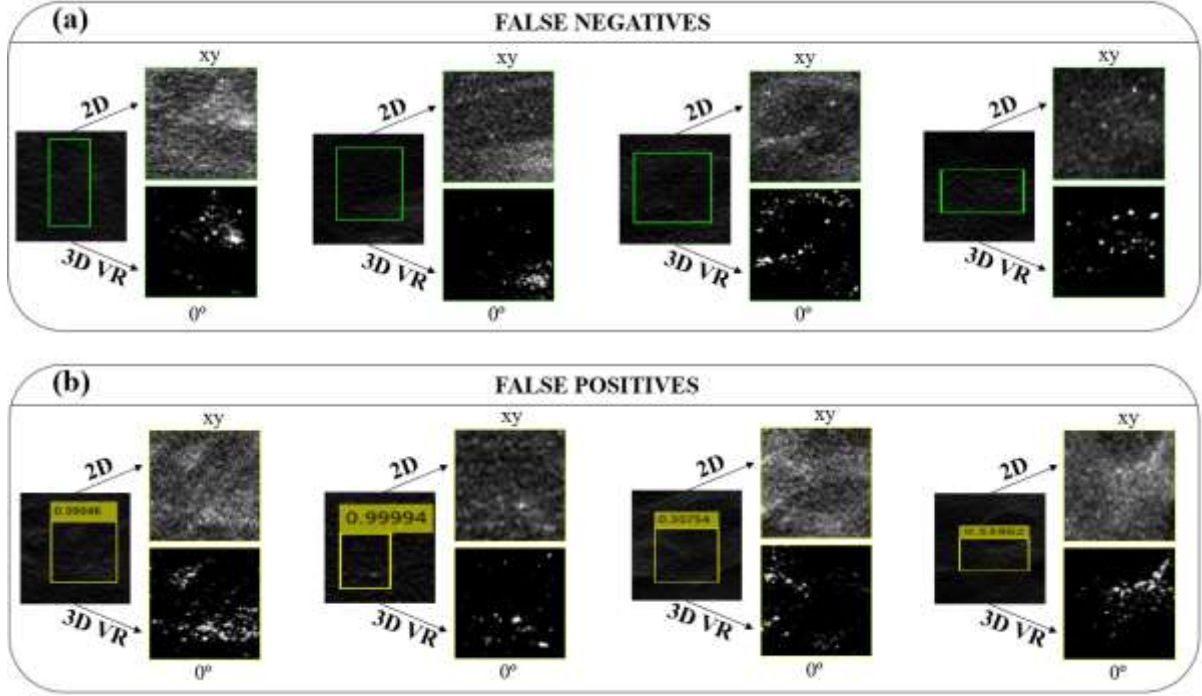


Figure 7.4. Example of four missed detections (false negatives) and four incorrect detections (FPs). The BB are visualized through 2D slice-by-slice in xy and 3D VR at 0° .

Visualization with 3D VR is very flexible and includes parameters that can significantly change its appearance, as is the case of transfer functions. The Figure 7.5 shows the displays of four detected MC obtained with 2D visualization and 3D VR using two different transfer functions.





Faster R-CNN detection (threshold: 0.90)	Visualization mode			
	2D Slice-by-slice	3D Volume rendering		
		Transfer Function 1	Transfer Function 2	
	xy	0°	0°	
	xz	90°	90°	
	xy	0°	0°	
	xz	90°	90°	
	xy	0°	0°	
	xz	90°	90°	
	xy	0°	0°	
	xz	90°	90°	
	xy	0°	0°	
	xz	90°	90°	
	xy	0°	0°	
	xz	90°	90°	
	xy	0°	0°	
	xz	90°	90°	
	xy	0°	0°	
	xz	90°	90°	

Figure 7.5. Example of four detection outputs. Green: Ground truth BB; yellow: predicted BB. The predicted results are visualized with 2D MIP slice-by-slice represented through xy and xz planes and 3D VR with two different angles (0° and 90°). Each 3D VR was obtained using two different transfer functions, allowing different levels of MC segmentations.

7.4. Discussion and conclusions

In this work, a Faster R-CNN detector was trained to detect MCs in DBT data and the preliminary results obtained were analyzed through two different forms of visualization: standard 2D slice-by-slice and 3D VR specifically developed for DBT. VR is presented as a supplementary visualization of the detected MCs, providing a more detailed and high quality complementary information.

A DBT dataset from the publicly available database at The Cancer Imaging Archive website [30] was used. The train dataset consisted in entire DBT images and also some regions of interest containing only the MCs. These smaller regions were included because the DBT images are much bigger than the ground truth boxes of MCs, reaching ratios of 30:1. As the size of the images was not changed in order keep the necessary spatial resolution to see the small microcalcifications, it was important to have training inputs with an emphasis on the object to be detected. Nevertheless, the test dataset only contains entire images, as happens in clinical or screening practice.

In this type of lesion detection task, the time required for the detector to give an answer about the input data is very important because it should be useful in real time clinical practice. 29 seconds to analyze a volume of DBT data (which can comprise ~130 million voxels) is reasonable but this value can be improved using computers with greater power. Also, this time is highly influenced by the feature extraction network. For this reason, in this preliminary work, we chose a network with a reasonable balance between time and accuracy (ResNet-18). However, other pre-trained networks that may show better results and different detection times should be studied.

The most used metric to analyze the performance of this type of detector is the FROC curve. The results obtained with this curve in Figure 7.2 and Table 7.3 reveal that it was possible to achieve a sensitivity of around 60% with 4 FP/image. These preliminary results are promising but need further improvement by adding more training data, optimizing some network parameters, training over a greater number of iterations and, as already mentioned, using different pre-trained CNNs.

The output results of Figure 7.3 were obtained by using a threshold of 0.9 (i.e., only scores above 0.9 were considered), which corresponds to a sensitivity of about 50% for 0.8 FP/image. Four examples of output from the Faster R-CNN were presented. The number of FP found for

this threshold (yellow BBs without a score) varies from three (in the first case) to zero (in the third and fourth case). The correctly detected MCs by the yellow BBs with scores that overlap the true BBs (green) were observed using the two mentioned visualization methods. In general, the MCs have a reasonable visibility in the xy plane with the 2D MIP slice-by-slice over five adjacent slices but are distorted in xz , losing some definition due to the larger voxel size in z . The 3D VR at 0° and 90° can be directly compared with the 2D visualization in the xy and xz planes, respectively. For all cases, there is better contrast and less noise in the VR at 0° , with better discernment of the MCs. This superior definition is noticeable when comparing the VR at 90° with the xz plane of the 2D visualization. In the VR at 90° there is a clear discrimination of the MCs, and it is possible to observe quite clearly the calcifications individually and with some degree of reality.

It is also important to analyze some situations where the detection was not correct (Figure 7.4). In the case of false negatives, there were prominent lesions that the algorithm did not detect (Figure 7.4 (a) last column) and others where the MCs were somehow masked, making their detection difficult (Figure 7.4 (a) third column). In the case of FPs, in fact, there were some situations where, even to the human eye, doubt could be raised (Figure 7.4 (b) second and third column). But, in the remaining situations, there is essentially a spiculated noise that was interpreted as MC. It is therefore important to further improve the quality of detection.

On the other hand, the flexibility of visualization using VR is demonstrated with the images in Figure 7.5. In addition to have the spatial distribution in the three directions (x , y and z), with different transfer functions we can filter the data to a greater or lesser extent and, thus, segment better some lesions, such as MCs. The transfer functions used in this work have the opacity/color on the y -axis and the intensity values on the x -axis. For intensities below a "A" value the object data is transparent, while intensity values above "B" ($A < B$) correspond to completely opaque voxels. Between A and B the opacity values follow a linear distribution. From transfer function 1 to transfer function 2 (Figure 7.5) the value of B has been increased to reduce the contribution to the visualization of objects with lower intensities, making those with higher intensities stand out, such as MCs. In this way, it was possible to obtain a "cleaner" visualization, as seen in Figure 7.5 in column of the transfer function 2. This rendering parameter is a great advantage in noisy data as can be seen in the last case of Figure 7.5.

During training, no distinction was made between the different types of breast density. However, different densities correspond to data with slightly different histograms. In the

detection/analysis step, it is important to understand if the detector behaves in the same way for different densities (for example, it is known that some lesions are more difficult to detect in dense breasts than in fat breasts). From the comparison made between the detection and visualization of the four density groups, we can infer that there were no differences between them.

As already mentioned, as far as we know, this is the first work about MCs detection and localization in a whole DBT image using deep learning CNNs such as Faster R-CNN. Of the few published works found in this area, all refer to soft tissue as masses. [26] developed a CAD system for the prescreening of ROIs and discrimination of true masses and FPs in DBT using a Faster R-CNN. For lesion-based mass detection, the sensitivity of their R-CNN based CAD was 90% at 1.54 FP/volume. Later, the same group, compared this work to a framework of a 3D-Mask R-CNN for mass detection and segmentation [27]. For lesion-based mass detection, the sensitivity of the 3D-Mask R-CNN based CAD (segmentation) was 90% with 0.8 FPs/lesion, whereas the sensitivity of the Faster R-CNN based CAD was 90% at 2.37 FPs/lesion. [23] developed a single-phase deep learning detection model for masses and architectural distortions and achieved a sensitivity of 65% at 2 FPs/breast. [28] propose a very interesting work on Faster R-CNN that uses mammary gland distribution as a prior information to detect architectural distortions in DBT and achieved a sensitivity of 80% at 1.85 FPs/volume for all architectural distortions types.

A fair and direct comparison between our results and these published data is not possible because they analyze completely different lesions, those are already optimized studies and of different characteristics (for example, some use ROIs and not the whole image to locate the lesions). Furthermore, although architectural distortions are quite difficult to locate, masses are more reasonable. Although masses have densities similar to the rest of the breast tissue and are often camouflaged, they are larger than microcalcifications, facilitating training and learning. It is possible to use images with less resolution and train more complex networks faster. Thus, we cannot make a comparison between our results and those already published, but we can conclude that, despite our high FP values in this preliminary study, there is potential to improve and achieve results similar to those of the masses.

In conclusion, taking into account the preliminary results presented, we conclude that detection and location of MCs in DBT can be automatically achieved using Faster R-CNN and visualization of these results can benefit from another approach such as 3D VR.

ACKNOWLEDGMENTS

This work was supported by Universidade de Lisboa (PhD grant) and Fundação para a Ciência e Tecnologia – Portugal (Grant No. SFRH/BD/135733/2018 and FCT-IBEB Strategic Project UIDB/00645/2020).

References

1. Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries*. CA Cancer J Clin, 2021. **71**(3): p. 209-249.
2. Tabár, L., et al., *The incidence of fatal breast cancer measures the increased effectiveness of therapy in women participating in mammography screening*. Cancer, 2019. **125**(4): p. 515-523.
3. Zackrisson, S., et al., *One-view breast tomosynthesis versus two-view mammography in the Malmö Breast Tomosynthesis Screening Trial (MBTST): a prospective, population-based, diagnostic accuracy study*. The Lancet Oncology, 2018. **19**(11): p. 1493-1503.
4. Bernardi, D., et al., *Breast cancer screening with tomosynthesis (3D mammography) with acquired or synthetic 2D mammography compared with 2D mammography alone (STORM-2): a population-based prospective study*. The Lancet Oncology, 2016. **17**(8): p. 1105-1113.
5. Lång, K., et al., *Performance of one-view breast tomosynthesis as a stand-alone breast cancer screening modality: results from the Malmö Breast Tomosynthesis Screening Trial, a population-based study*. Eur Radiol, 2016. **26**(1): p. 184-190.
6. Gilbert, F.J., et al., *Accuracy of Digital Breast Tomosynthesis for Depicting Breast Cancer Subgroups in a UK Retrospective Reading Study (TOMMY Trial)*. Radiology, 2015. **277**(3): p. 697-706.
7. Hofvind, S., et al., *Digital Breast Tomosynthesis and Synthetic 2D Mammography versus Digital Mammography: Evaluation in a Population-based Screening Program*. Radiology, 2018. **287**(3): p. 787-794.
8. Freer, P.E., et al., *Clinical implementation of synthesized mammography with digital breast tomosynthesis in a routine clinical practice*. Breast Cancer Research and Treatment, 2017. **166**(2): p. 501-509.
9. Food and Drug Administration (FDA) U.S. . *Premarket Approval application supplement for the Selenia Dimensions 3D System*. 2013 [cited 2021 May].
10. Good, W.F., et al., *Digital breast tomosynthesis: a pilot observer study*. AJR Am J Roentgenol, 2008. **190**(4): p. 865-9.
11. Gur, D., et al., *Digital breast tomosynthesis: observer performance study*. AJR Am J Roentgenol, 2009. **193**(2): p. 586-91.
12. Caumo, F., et al., *Digital Breast Tomosynthesis with Synthesized Two-Dimensional Images versus Full-Field Digital Mammography for Population Screening: Outcomes from the Verona Screening Program*. Radiology, 2018. **287**(1): p. 37-46.
13. Fenton, J.J., et al., *Influence of Computer-Aided Detection on Performance of Screening Mammography*. New England Journal of Medicine, 2007. **356**(14): p. 1399-1409.
14. Lehman, C.D., et al., *Diagnostic Accuracy of Digital Screening Mammography With and Without Computer-Aided Detection*. JAMA Internal Medicine, 2015. **175**(11): p. 1828-1837.
15. Katzen, J. and K. Dodelzon, *A review of computer aided detection in mammography*. Clinical Imaging, 2018. **52**: p. 305-309.
16. Sechopoulos, I., J. Teuwen, and R. Mann, *Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art*. Seminars in Cancer Biology, 2020.
17. Venson, J.E., et al., *A Case-Based Study with Radiologists Performing Diagnosis Tasks in Virtual Reality*. Stud Health Technol Inform., 2017. **245**: p. 244-248.

18. Suetens, P., *Medical image analysis*, in *Fundamentals of Medical Imaging* 2009, Cambridge University Press: New York. p. 159-189.
19. Girshick, R., et al. *Rich feature hierarchies for accurate object detection and semantic segmentation*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014.
20. Girshick, R. *Fast R-CNN*. in *Proceedings of the IEEE international conference on computer vision*. 2015.
21. Ren, S., et al., *Faster R-CNN: Towards real-time object detection with region proposal networks*. *Advances in neural information processing systems*, 2015. **28**: p. 91-99.
22. Lai, X., W. Yang, and R. Li, *DBT Masses Automatic Segmentation Using U-Net Neural Networks*. *Comput Math Methods Med*, 2020. **2020**: p. 7156165.
23. Buda, M., et al. *Detection of masses and architectural distortions in digital breast tomosynthesis: a publicly available dataset of 5,060 patients and a deep learning model*. 2020. arXiv:2011.07995.
24. Fotin, S., et al., *Detection of soft tissue densities from digital breast tomosynthesis: comparison of conventional and deep learning approaches*. *SPIE Medical Imaging*. Vol. 9785. 2016: SPIE.
25. Samala, R.K., et al., *Mass detection in digital breast tomosynthesis: Deep convolutional neural network with transfer learning from mammography*. *Med Phys*, 2016. **43**(12): p. 6654-6654.
26. Fan, M., et al., *Computer-aided detection of mass in digital breast tomosynthesis using a faster region-based convolutional neural network*. *Methods*, 2019. **166**: p. 103-111.
27. Fan, M., et al., *Mass Detection and Segmentation in Digital Breast Tomosynthesis Using 3D-Mask Region-Based Convolutional Neural Network: A Comparative Analysis*. *Frontiers in Molecular Biosciences*, 2020. **7**(340).
28. Li, Y., et al., *Deep learning of mammary gland distribution for architectural distortion detection in digital breast tomosynthesis*. *Physics in Medicine & Biology*, 2021. **66**(3): p. 035028.
29. Badano, A., et al., *Evaluation of Digital Breast Tomosynthesis as Replacement of Full-Field Digital Mammography Using an In Silico Imaging Trial*. *JAMA Network Open*, 2018. **1**(7): p. e185474-e185474.
30. VICTRE. *The VICTRE Trial: Open-Source, In-Silico Clinical Trial For Evaluating Digital Breast Tomosynthesis*. 2018 [cited 2021; Available from: <https://wiki.cancerimagingarchive.net/display/Public/The+VICTRE+Trial%3A+Open-Source%2C+In-Silico+Clinical+Trial+For+Evaluating+Digital+Breast+Tomosynthesis>].
31. ImageNet. 2021 [cited 2021 October]; Available from: <http://www.image-net.org>.
32. MathWorks. *MATLAB trainFasterRCNNObjectDetector*. 2021 [cited 2021 October]; Available from: <https://www.mathworks.com/help/vision/ref/trainfasterrcnnobjectdetector.html>.
33. Bunch, P., et al., *A Free Response Approach To The Measurement And Characterization Of Radiographic Observer Performance*. *Application of Optical Instrumentation in Medicine VI*. Vol. 0127. 1977: SPIE.
34. VTK. *Visualization Toolkit - VTK*. 2020 [cited 2020 February]; Available from: <http://www.vtk.org/>.
35. Schroeder, W., K. Martin, and B. Lorensen, *The Visualization Toolkit: An Object-oriented Approach to 3D Graphics*. 4rd ed 2006, USA: Kitware.
36. Mota, A.M., et al. *Calculation of transfer functions for volume rendering of breast tomosynthesis imaging*. in *15th International Workshop on Breast Imaging (IWBI2020)*. 2020. Leuven, Belgium: SPIE.

Final considerations

8

Breast cancer is one example where progresses in technology and therapy have made the difference in the decline of mortality. Despite being a relatively new technique, DBT has already consolidated its role in breast cancer imaging. With DBT, it is possible to analyze the acquired breast tissue over several slices and, thus, better separate the depth information. The work presented in this thesis focused on the development of computational methodology to increase the diagnostic value of DBT, namely in terms of visualization and automatic detection of lesions. The main objective of this thesis, which was to directly respond to the current challenges in the DBT data visualization and detection of lesions in DBT, was successfully achieved. With proper developed software and new processing techniques, it was possible to render the DBT volume data with quality, allowing a complementary visualization of DBT exams. Furthermore, this work achieved very positive results with regard to the classification of DBT images for the presence or absence of MCs. And, going a little further, through a preliminary but very interesting development, the exact position of these MCs can be determined and the detected MCs rendered.

Currently, DBT data visualization is done with a slice-by-slice mode or in a continuous loop where radiologists explore and evaluate each exam. Such a procedure leads to longer inspection (being a disadvantage in a screening environment) and requires a mental 3D reconstruction of the tissues looking at adjacent slices (complicating the interpretation of the morphology of lesions spread over several slices). In this way, a complementary mode of visualizing DBT data through 3D VR was studied and developed.

Chapters 2, 3, 4 and 5 of this thesis were specifically dedicated to visualization. In chapters 2 and 3 some critical rendering parameters were explored and optimized. In chapters 4 and 5 methodologies to improve the quality of rendered images were developed.

In chapter 2, the work was developed to answer to one of the biggest challenges of VR: finding transfer functions that result in adequate DBT data visualizations. As both the intensity and the gradient magnitude histograms showed similar shapes among the diverse clinical cases observed, it was possible to find a linear relation between these values and the opacity/color. In this way, a consistent and reproducible methodology for the automatic determination of transfer functions was presented, allowing different 3D VR visualizations for DBT data. This work provides the tools for the automatic generation of transfer functions directly from the data (avoiding the trial and error procedure that often leads to exhaustion or incorrect visualization)

but a wide validation in clinical environment for radiologists to evaluate and select the transfer functions that result in a real clinical utility of VR visualization is missing.

Additionally, two parameters that also affect image quality were studied in chapter 3: voxel size in z direction and sampling distance. As the reconstructed DBT data has anisotropic voxels, although the VR at 0° shows a good quality, when a rotation around the object is carried out, there is a distortion and loss of quality. In this way, the DBT reconstructed voxels were made isotropic using different interpolation functions from the VTK library, changing their size from $0.085 \times 0.085 \times 1.0 \text{ mm}^3$ to $0.085 \times 0.085 \times 0.085 \text{ mm}^3$. In addition, two other interpolation parameters that influence the results (window half width and blur factor in z direction) were evaluated. About the study of sampling distance (distance between discrete points of accumulation of intensity values weighted by transparency or opacity along the ray), it was observed that the quality of the VR and the time required for the rendering were significantly affected by it. In this case, an effort was made to find appropriate sampling distance values that allowed a balance between good image quality at all angles and the time required for rendering. With the exhaustive study of the interpolation functions and their parameters, in conjunction with the appropriate sampling distance value, a significant improvement in the quality of the VR of DBT data was achieved with this work. Furthermore, the results obtained with the phantom were consolidated through the analysis of the VR of a real clinical case.

In addition to the aforementioned parameters, new image processing techniques that improve the rendering in two points were developed: (1) noise regularization and (2) reduction of the “out-of-plane” artifact. For the first point, two 3D algorithms to minimize the total variation (TV) of the data were implemented. The algorithms differ in how they are applied to the volume. One considers a slice-by-slice optimization and the other considers the intensity values of adjacent slices to make this optimization on each voxel. The application of this type of filters has shown excellent results in decreasing noise while preserving the edges of objects and it is a straightforward not time-consuming approach. In addition, the analysis of the results was done using the visualization obtained through 3D VR, allowing a better evaluation of the effects in all directions. The results were very encouraging with a significantly increased CNR at 0° and 90° . On the other hand, as the out-of-plane artifact is a current drawback in DBT imaging, for the second point, a methodology to study the application of blind deconvolution to DBT data was proposed. A visual decrease of the artifact, demonstrated by a numerical reduction of about 23% was observed. In this way, both quantitative and qualitative results showed the relevance

of this approach in improving the image quality in DBT by reducing the out-of-plane artifact without introducing noise or other artifacts after deconvolution.

In recent years, the increase in computational power and the growth of available large databases have allowed the development of deep learning AI algorithms composed of multilayer CNNs. These algorithms have emerged as a potential solution for the automated detection of breast cancer. Chapters 6 and 7 were specifically dedicated to the automatic classification of DBT images and detection of MCs. Chapter 6 described the work developed for the automatic classification of DBT for the presence or absence of MCs. On the other hand, Chapter 7 referred to the detection and localization of these MCs and posterior visualization of them using the VR techniques described above.

In chapter 6, four popular deep CNNs were trained using data from a public DBT database and compared with one new architecture proposed in this thesis. The main task of these trainings was to receive the entire DBT image and output its classification by absence or presence of MCs, without prior identification of any regions. With this comprehensive study, very promising results in detecting MCs were achieved with a maximum AUC value of 94.19% for the Googlenet and 91.17% for the new implemented network. Our CNN also had the particularity of being the fastest, thus becoming a very interesting model to be considered in the future. The detection of MCs with each CNN was also greatly influenced by the type of input data, differing between the different pre-processing methodologies implemented and emphasizing the fact that attention should be paid to the quality of the data when CNNs are trained from scratch.

The work in chapter 7 results from the connection between the knowledge developed for AI algorithms and VR methodologies. The main objective of this chapter was to create a conductive pipeline that contemplated the automatic location of MCs and their subsequent visualization through 3D VR. The work developed here was not optimized for detection and could be greatly improved in the future. A faster R-CNN detector was trained to detect and locate MCs in DBT data and the obtained results were analyzed through the FROC curve. It was possible to achieve a sensitivity of around 60% with 4 FP/image. As mentioned, these are preliminary results that are very promising but should be further improved by adding more training data, optimizing some network parameters, training over a greater number of iterations and using different pre-trained feature extraction networks (the ResNet-18 was used). After the detection, the MCs were analyzed through standard 2D visualization and the 3D VR specifically

developed for DBT data. The advantage of a 3D visualization with VR for the MCs is quite clear in this work due to the possibility of rotating around the MCs and having an immediate notion of its morphology in space. In addition, in situations with more noise involving the MCs, the use of different transfer functions with different characteristics, allowed a much clearer discernment of the MCs detected by the network. By slightly changing the transfer functions, the data were filtered to a greater or lesser extent and thus different rendered images of the MCs were obtained.

This part of the work (chapter 6 and 7) focused on these types of lesions since MCs are one of the most important biomarkers for breast cancer, especially in cases of non-palpable lesions. Furthermore, due to its small size (which requires high spatial resolution, and thus the need for high computing power), until recently most of the work developed with AI in DBT referred to soft tissue lesions, such as masses or architectural distortions.

The biggest limitation of the work developed in this thesis in terms of visualization is the lack of clinical evidence through a comprehensive clinical study to evaluate the complementarity of visualization of DBT data with VR. The procedure for this study is already written but it is in the process of approval.

On the other hand, considering the training of the deep learning CNNs to detect and classify lesions, the major limitation was the database and the high computational power needed. The VICTRE database was chosen because it is public (and thus allows a fair comparison of new methods) but it consists of simulated realistic DBT data. In this way, to generalize the accuracy of the developed methods in a correct way, it is important to use a real database and data from different manufacturers, with different image characteristics.

As future work, what was mentioned as limitation should be considered in order to, by surpassing them, support and consolidate all the developed work. In addition, there are some specific tasks that can be considered in the future to improve the work presented:

- Study the impact of transfer functions on the specific visualization of soft tissues (such as masses or architectural distortions);
- Analyze the impact of the image processing methodologies on VR at different angles others than 0° and 90°;
- Improve the PSF estimation in the blind deconvolution and test other deconvolution approaches;

- In the training of deep CNNs, include more data, consider 3D architectures, diversify the lesions and include data obtained from other views (not only craniocaudal) and reconstruction algorithms.

With the work carried out, this thesis has contributed to increasing the clinical value of this new breast imaging technique, DBT, in the detection and improved diagnosis of breast cancer. Although there is still work to be done, namely in the clinical validation by specialists and in the optimization of the developed tools, this thesis achieved innovative and consistent results in the visualization and detection of lesions in DBT.

