UNIVERSIDADE DE SÃO PAULO–USP
ESCOLA DE ENGENHARIA DE SÃO CARLOS
DEPARTAMENTO DE ENGENHARIA ELÉTRICA E DE COMPUTAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Helder Cesar Rodrigues de Oliveira

**Novas Abordagens para a Detecção
Automática da Distorção Arquitetural
na Mamografia Digital e Tomossíntese
Mamária**

Helder Cesar Rodrigues de Oliveira

**Novas Abordagens para a Detecção
Automática da Distorção Arquitetural
na Mamografia Digital e Tomossíntese
Mamária**

Tese de doutorado apresentada à Escola de Engenharia de São Carlos da Universidade de São Paulo como parte dos requisitos para a obtenção do título de Doutor em Ciências, Programa de Engenharia Elétrica.

Área de concentração: Processamento de Sinais e Instrumentação

Orientador: Prof. Dr. Marcelo Andrade da Costa Vieira

São Carlos

2019

Trata-se da versão corrigida da tese. A versão original se encontra disponível na EESC/USP que aloja o Programa de Pós-Graduação de Engenharia Elétrica.

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha catalográfica elaborada pela Biblioteca Prof. Dr. Sérgio Rodrigues Fontes da
EESC/USP com os dados inseridos pelo(a) autor(a).

R474n   Rodrigues de Oliveira, Helder Cesar
Novas Abordagens para a Detecção Automática da
Distorção Arquitetural na Mamografia Digital e
Tomossíntese Mamária / Helder Cesar Rodrigues de
Oliveira; orientador Marcelo Andrade da Costa Vieira.
São Carlos, 2019.

Tese (Doutorado) - Programa de Pós-Graduação em
Engenharia Elétrica e Área de Concentração em
Processamento de Sinais e Instrumentação -- Escola de
Engenharia de São Carlos da Universidade de São Paulo,
2019.

1. Mamografia Digital. 2. Filtro de Gabor. 3.
Tomossíntese Digital Mamária. 4. Distorção
Arquitetural. I. Título.

FOLHA DE JULGAMENTO

Candidato: Bacharel **HELDER CESAR RODRIGUES DE OLIVEIRA.**

Título da tese: "Novas abordagens para detecção automática de distorção arquitetural na mamografia digital e tomossíntese mamária".

Data da defesa: 26/08/2019

Comissão Julgadora:

Resultado:

Prof. Associado **Marcelo Andrade da Costa Vieira**
(Orientador)

(Escola de Engenharia de São Carlos/EESC)

APROVADO

Prof. Dr. **Ricardo José Ferrari**

(Universidade Federal de São Carlos/UFSCar)

APROVADO

Prof. Dr. **Marco Antônio Piteri**

(Universidade Estadual Paulista "Júlio de Mesquita Filho"/UNESP – Presidente Prudente)

APROVADO

Profa. Titular **Agma Juci Machado Traina**

(Instituto de Ciências Matemáticas e de Computação/ICMC-USP)

APROVADO

Prof. Associado **Adilson Gonzaga**

(Escola de Engenharia de São Carlos/EESC)

APROVADO

Coordenador do Programa de Pós-Graduação em Engenharia Elétrica:

Prof. Associado **Ivan Nunes da Silva**

Presidente da Comissão de Pós-Graduação:

Prof. Titular **Murilo Araujo Romero**

*Dedico este trabalho aos meus pais, Lúcia e Elmo, e a minha esposa Amanda.
Sem o apoio deles nada disso teria sido possível.*

Agradecimentos

Este trabalho não seria possível se não fosse pelo grande apoio de algumas pessoas que fizeram e fazem parte de minha vida, contribuindo de maneira significativa para que eu conseguisse alcançar meus objetivos.

Primeiramente gostaria de agradecer aos meu pais, Lúcia e Elmo, por TUDO: o apoio nos momentos difíceis, o incentivo de continuar os estudos e seguir os meus sonhos, além da compreensão quando não pude ir para casa quando estava muito atarefado. Sei o esforço e todo o sacrifício que vocês fizeram para que eu pudesse chegar até aqui. À minha esposa Amanda, minha companheira há mais de 10 anos que sempre me ouve, questiona e aconselha. É especialmente à vocês que dedico este trabalho de todo o meu coração. Sem vocês nada disso teria sido possível.

Ao Prof. Dr. Marcelo Andrade da C. Vieira, meu orientador durante essa jornada. Agradeço imensamente por toda a confiança investida desde o primeiro momento, todas as conversas e aconselhamentos, a atenção e paciência com que tem me atendido nos momentos que o procurei.

Aos professores Arianna Mencattini e Eugênio Martinelli, da *Università degli Studi di Roma "Tor Vergata"*, por terem me recebido para o doutorado sanduíche financiado pelo programa Erasmus Mundus (BEMUNDUS). Por todos os aconselhamentos, pessoais e profissionais, e as dúvidas tiradas nos momentos em que o trabalho parava.

Ao professor Rangaraj M. Rangayyan, por ter me recebido para conversar e ter apoiado a proposta deste trabalho quando era apenas uma ideia, além de ter me indicado a professora Arianna para o período sanduíche. Também não posso deixar de agradecê-lo pela indicação do pós-doutorado na *University of Calgary*.

Aos doutores Nestor de Barros, Juliana Catani, ambos do InRAD e ao doutor Carlos Erbolato da clínica Eco & Mama, por toda atenção durante a coleta das imagens, as dúvidas sanadas e as imagens marcadas. Toda a atenção e dedicação com que me atenderam foi essencial para que eu pudesse concluir esse trabalho.

Por último, mas não menos importante, agradeço a Deus, por ter me dado força nos momentos de dificuldade e todas as vezes que olhei para trás, além de ter olhado por aqueles que direta, ou indiretamente, causaram algum impacto em minha vida e neste trabalho.

*“All knowledge which ends in words
will die as quickly as it came to life,
with the exception of the written word:
which is its mechanical part.”*

(Leonardo da Vinci, Thoughts on Art and Life)

Resumo

Rodrigues de Oliveira, Helder Cesar **Novas Abordagens para a Detecção Automática da Distorção Arquitetural na Mamografia Digital e Tomossíntese Mamária.** 156 p. Tese de doutorado – Escola de Engenharia de São Carlos, Universidade de São Paulo, 2019.

O câncer de mama é a doença que mais acomete as mulheres em todo o mundo, sendo o tratamento mais eficaz se for diagnosticada em estágio inicial. A partir de 2011, nos programas de rastreamento de países desenvolvidos, vem sendo empregada uma nova modalidade de exame, a tomossíntese digital mamária (*Digital Breast Tomosynthesis* - DBT), que possui diversas vantagens se comparada à mamografia digital. No exame, o médico radiologista busca por sinais suspeitos na imagem, como: nódulos, microcalcificações e distorção arquitetural mamária (DAM). Sendo que, este último pode representar o estágio mais inicial de um câncer em formação, podendo se manifestar antes da formação de qualquer outra lesão. No entanto, a DAM é difícil de ser detectada pois modifica o tecido mamário de forma sutil, não havendo qualquer formação de massa ou a borda definida. Os sistemas computacionais de auxílio ao diagnóstico (*Computer-Aided Detection* - CAD) vêm apresentando alto desempenho na detecção de nódulos e microcalcificações mamárias, mas para o caso da DAM, o desempenho ainda é insatisfatório. Algumas limitações são normalmente reportadas nos algoritmos adotados para detectar automaticamente a DAM. O presente trabalho tem por objetivo propor novas abordagens para aumentar a precisão dos métodos computacionais de detecção: o uso de descritores de micro-padrões local para discriminação de áreas suspeitas; redução de falsos-positivos; uso do volume 3D fornecido pelo exame de DBT e; uso de arquitetura de aprendizagem profunda para discriminação e classificação de regiões suspeitas. Os diversos testes efetuados em cada proposta mostraram que é possível melhorar as taxas de detecção da DAM, mesmo para imagens de DBT onde ainda não há um esquema computacional de detecção bem estabelecido.

Palavras-chave: Mamografia Digital. Filtro de Gabor. Tomossíntese Digital Mamária. Distorção Arquitetural.

Abstract

Rodrigues de Oliveira, Helder Cesar **New Approaches for Automatic Detection of Architectural Distortion in Digital Mammography and Digital Breast Tomosynthesis**. 156 p. Ph.D. Thesis – São Carlos School of Engineering, University of São Paulo, 2019.

Breast cancer is the disease that most affects women worldwide and is the most effective treatment if it is diagnosed at early stages. Since 2011, in developed countries screening programs has been employed a new exam, the digital breast tomosynthesis (DBT), which has several advantages compared to the digital mammography. In the exam, the radiologist looks for suspicious signs in the image such as masses, microcalcifications and architectural distortion of breast (ADB). Since the later may represent the earliest sign of a cancer in formation, it may manifests before the formation of any other lesion. However, ADB is difficult to be detected due to its subtly changes the breast tissue, with no mass or defined shape. Computer-aided detection (CAD) systems have shown good results in the detection of masses and microcalcifications, however, for ADB the performance is still poor. Several weakness are reported in the pipeline adopted to automatic detection of ADB. The present work aims to propose new approaches to increase the accuracy of the current CAD pipeline: the use of local micro-pattern descriptors to discriminate suspicious areas; false-positives reduction; automatic detection of ADB in DBT images using and tridimensionality of the exam and; use of deep learning architecture to discriminate and classify suspicious regions. The several tests performed on each proposal showed that it is possible to improve the detection rates even for DBT images where there is no established CAD pipeline.

Keywords: Digital Mammography. Gabor Filter. Digital Breast Tomosynthesis. Architectural Distortion.

Lista de ilustrações

Figura 1	Mamografia com uma distorção arquitetural marcada por radiologista. O destaque apresenta a DAM em maiores detalhes.	24
Figura 2	Ilustração de exame mamográfico convencional, a Mamografia Digital de Campo Total (FFDM).	30
Figura 3	Ilustração da ocorrência da sobreposição de tecidos em mamografia 2D.	31
Figura 4	Ocorrência da sobreposição de tecidos em uma mamografia hipotética.	31
Figura 5	Ilustração de exame de tomossíntese digital mamária (DBT).	32
Figura 6	Sequência de fatias mamográficas produzidas em um exame de DBT, no qual o problema da sobreposição de tecidos é solucionado.	33
Figura 7	Mamografica com marcações referente às regiões suspeitas encontradas pelo CAD ImageChecker ® v.10.0.	34
Figura 8	Sequência de passos usualmente aplicados para a detecção da distorção arquitetural mamária.	37
Figura 9	Exemplo de mamografias nas incidências CC e MLO.	39
Figura 10	Filtro de Gabor em vista lateral e superior.	40
Figura 11	Exemplo de mamografias segmentadas na região que compreende a mama.	41
Figura 12	Mamografias com a região da mama segmentada.	42
Figura 13	Resultado do realce da estrutura mamária.	43
Figura 14	Resultado após a segmentação e rastreo de regiões suspeitas.	45
Figura 15	Exemplo da ordenação e corte da lista de regiões suspeitas.	46
Figura 16	Vizinhança de um pixel arbitrário g_0 considerando raio unitário.	47
Figura 17	Exemplo de aplicação do LBP.	48
Figura 18	Histograma de códigos LBP.	49
Figura 19	Gráfico da função <i>Sigmoid</i> com diferentes parâmetros β	51
Figura 20	Neurônio artificial <i>Perceptron</i>	52
Figura 21	Imagens mamográficas do banco DDSM.	58

Figura 22	Imagens clínicas de mamografia FFDM (2D) do banco de imagens do InRAD.	59
Figura 23	Imagens clínicas de DBT (fatias).	60
Figura 24	Imagens clínicas após a marcação do radiologista.	61
Figura 25	Localização da proposta #1 dentro do <i>pipeline</i> para a detecção automática da DAM.	62
Figura 26	Localização da proposta #2 dentro do <i>pipeline</i> para a detecção automática da DAM.	63
Figura 27	Proposta para o descarte eficiente de falsos-positivos.	64
Figura 28	Localização da proposta #3 dentro do <i>pipeline</i> para a detecção automática da DAM.	65
Figura 29	Proposta de um <i>pipeline</i> para a detecção de DAM em imagens de DBT.	66
Figura 30	Exemplo de trilha (em vermelho) montada a partir da conexão de diferentes regiões suspeitas em cada fatia.	67
Figura 31	Localização da proposta #4 dentro do <i>pipeline</i> para a detecção automática da DAM.	68
Figura 32	<i>Autoencoder</i> proposto para discriminação de regiões com DAM.	69
Figura 33	Modelo <i>encoder</i> , retirado do <i>autoencoder</i> apresentado na Figura 32, conectado a uma rede neural.	69
Figura 34	Curvas ROC calculadas utilizando os descritores LMP, LBP e Haralick-4.	72
Figura 35	Curvas ROC resultante da discriminação de DAM e tecido normal com os descritores Energia, LMP e LBP, em imagens mamografia digital.	74
Figura 36	Representação das taxas de erros de classificação dos descritores LMP, LBP e Haralick-4.	75
Figura 37	Taxa de sucesso com base no processo usual, que utiliza o cálculo da média na imagem de magnitudes do filtro de Gabor, para análise e remoção de FPs.	76
Figura 38	Taxa de sucesso de outras 11 métricas estatísticas usando ainda a imagem de magnitudes de Gabor.	77
Figura 39	Taxa de sucesso de outras 11 métricas estatísticas usando o mapa de Texturas gerado pelo LMP.	78
Figura 40	Curva <i>Free</i> ROC referente ao processo de classificação das trilhas a nível de ROIs	79
Figura 41	Exame de DBT na incidência MLO, com as respectivas trilhas após a classificação sendo uma FP, na qual a DAM é representada por duas trilhas distintas.	80
Figura 42	Exame de DBT na incidência MLO, com uma trilha (em azul claro) muito próxima da DAM.	80

Figura 43	Exame de DBT na incidência CC, com as respectivas trilhas após a classificação sendo uma FP.	81
Figura 44	Curvas ROC referente aos <i>datasets</i> utilizados para treinamento e predição.	82

Lista de siglas

AUC *Area Under the Curve* - Área Sob a Curva

CAD *Computer-Aided Detection* - Detecção Assistida por Computador

CC Craniocaudal

DAM Distorção Arquitetural Mamária

DBT *Digital Breast Tomosynthesis* - Tomossíntese Digital Mamária

DDSM *Digital Database for Screening Mammography*

FFDM *Full Field Digital Mammography* - Mammografia Digital de Campo Total

INCA Instituto Nacional de Câncer

InRAD Instituto de Radiologia da Faculdade de Medicina da USP

LBP *Local Binary Pattern*

LMP *Local Mapped Pattern*

MLO Mediolateral Oblíqua

ROC *Receiver Operating Characteristic*

ROI *Region of Interest* - Região de Interesse

SVM *Support Vector Machine*

Sumário

1	Introdução	23
1.1	Objetivos	27
1.2	Organização da Tese	27
2	O Câncer de Mama e o Rastreamento	29
2.1	Mamografia Digital	30
2.2	Tomossíntese Digital Mamária	32
2.3	Sistemas de Auxílio ao Diagnóstico	33
3	Deteccção Automática da Distorção Arquitetural na Mamografia	37
3.1	Segmentação da Mama	38
3.2	Análise da Estrutura Mamária	42
3.3	Rastreamento de Regiões Suspeitas	43
3.4	Redução de Falsos-Positivos	45
3.5	Extração de Características	46
3.5.1	Descritores Baseados em Micro-padrões Local	47
3.6	Classificação	51
3.7	Trabalhos Correlatos	53
4	Materiais e Métodos	57
4.1	Bancos de Imagens Mamográficas Clínicas	57
4.1.1	<i>Digital Database for Screening Mammography</i>	57

4.1.2	Imagens Digitais	58
4.2	Métodos Propostos	61
4.2.1	Uso de Descritores de Micro-Padrões Local para a Discriminação de Regiões com Distorção Arquitetural	62
4.2.2	Redução de Falsos-Positivos	63
4.2.3	Detecção de Distorção Arquitetural em Imagens de Tomossíntese Mamária	64
4.2.4	Uso de Arquitetura de Aprendizagem Profunda para o Reconheci- mento de Áreas com DAM	67
5	Resultados e Discussão	71
5.1	Uso de Descritores de Micro-Padrões Local para a Discriminação de Regiões com Distorção Arquitetural	71
5.2	Redução de Falsos-Positivos	76
5.3	Detecção de Distorção Arquitetural em Imagens de Tomossíntese Mamária	78
5.4	Uso de Arquitetura de Aprendizagem Profunda para o Reconhecimento de Áreas com DAM	81
6	Conclusões e Trabalhos Futuros	85
7	Publicações Oriundas da Tese	89
7.1	Artigo 1: XII Workshop de Visão Computacional	91
7.2	Artigo 2: <i>SPIE Medical Imaging 2017</i>	99
7.3	Artigo 3: <i>SPIE Medical Imaging 2018</i>	111
7.4	Artigo 4: <i>Biomedical Signal Processing and Control</i>	123
7.5	Artigo 5: <i>SPIE Medical Imaging 2019</i>	137
	Referências	149

Introdução

O câncer de mama é a doença que mais ameaça a população feminina em nível mundial. Devido a inexistência de um tratamento que garanta a sua cura, as chances de sobrevivência do paciente aumentam consideravelmente se a doença for detectada em estágio inicial e o tratamento adequado for aplicado (WHO, 2019; INCA, 2019).

Para a descoberta prematura do que pode vir a ser um câncer de mama, a mamografia é o exame mais utilizado em todo o mundo (KARELLAS; VEDANTHAM, 2008; GLYNN et al., 2011). Porém, ela apresenta uma grande limitação: a sobreposição de tecidos e estruturas mamárias na imagem, que ocorre devido a projeção de um objeto tridimensional nesse caso a mama do paciente, em um plano bidimensional, o detector e consequentemente a imagem.

Um exame contemporâneo a mamografia é a Tomossíntese Digital da Mama (*Digital Breast Tomosynthesis* - DBT) também chamada de "mamografia 3D", que será explicada em detalhes na Seção 2.2. Em essência, a DBT é como uma tomografia computadorizada da mama, na qual durante o processo de aquisição o tubo de raios-X se move em torno da mama. Se comparada com a mamografia digital, a DBT permite uma maior precisão no diagnóstico pois o conjunto de imagens gerados são como uma imagem para cada seção da mama (RAFFERTY et al., 2013; RAY et al., 2015; DURAND et al., 2015; VEDANTHAM et al., 2015).

Durante a análise de uma imagem mamográfica, o radiologista tem a sua atenção voltada à busca de nódulos e microcalcificações mamárias, pois essas podem apontar a presença de um câncer. No entanto, recentemente os médicos radiologistas e a comunidade científica têm se voltado para um outro sinal indicativo de câncer de mama, a Distorção Arquitetural Mamária (DAM).

Ao contrário das outras anomalias, a DAM é o sinal mais precoce de um câncer em formação, podendo surgir antes mesmo de qualquer outra anomalia. Além disso, se a DAM for detectada precocemente, o prognóstico no tratamento do câncer é muito superior

do que no caso da detecção precoce de microcalcificações e nódulos (GAUR et al., 2013; BAHL et al., 2015).

Figura 1: Mamografia com uma distorção arquitetural marcada por radiologista. O destaque apresenta a DAM em maiores detalhes.

Fonte: O autor (2019).

A DAM é caracterizada pela contração sutil do tecido mamário, não possui borda bem definida e não ocorre a formação de massa visível, desse modo a paciente não é capaz de detectá-la com o autoexame. A textura do tecido mamário se modifica de forma tão sutil, que muitas vezes é impossível de ser detectada pelo olho humano (RAY et al., 2015).

Devido à dificuldade de localização, a taxa média de detecção da DAM pelos radiologistas é inferior a 50%, se tornando então o achado mais comum em exames retrospectivos de mamografia que apresentaram casos de falso-negativo (RANGAYYAN; BANIK; DESAUTELS, 2010; GAUR et al., 2013; BAHL et al., 2015). Um exemplo de mamografia

com DAM, marcada por radiologista, é apresentado na Figura 1.

Com o objetivo de auxiliar os radiologistas no diagnóstico do câncer de mama, sistemas computacionais de auxílio ao diagnóstico (*Computer-Aided Detection* - CAD) são normalmente utilizados na prática clínica. Esses softwares analisam as imagens do exame e apontam de forma automática regiões suspeitas, fornecendo uma “segunda opinião” ao médico radiologista (DEAN; ILVENTO, 2006; DOI, 2007; YANG et al., 2007).

Quando assistidos por tal tipo de sistema, os radiologistas são capazes de detectar, na mamografia digital, aproximadamente 90% das microcalcificações e nódulos. No entanto, para a detecção da DAM o desempenho dos sistemas CAD é insatisfatório pois, quando a detectam, normalmente acontece devido a sua capacidade de detecção de nódulos espiculados, que correspondem a uma outra classe de lesão (BAKER et al., 2003; RANGAYYAN; BANIK; DESAUTELS, 2010; SULEIMAN et al., 2016; RUSSO, 2018; KOHLI; JHA, 2018).

A revisão da literatura mostrou que a sequência de métodos (*pipeline*) que tem sido adotada para a detecção automática da DAM em imagens mamográficas consiste essencialmente de 6 etapas distintas (RANGAYYAN; AYRES, 2006; BOZEK; DELAC; GRGIC, 2008; TANG et al., 2009; JASIONOWSKA; PRZELASKOWSKI, 2010; BANIK; RANGAYYAN; DESAUTELS, 2011; BISWAS; MUKHERJEE, 2011; ANAND; RATHANA, 2013; LAKSHMANAN et al., 2014; NETPRASAT; AUEPHANWIRIYAKUL; THEERA-UMPON, 2014; KAMRA et al., 2016; ZYOUT; TOGNERI, 2015; LIU et al., 2016a; LIU et al., 2016b; ZYOUT; TOGNERI, 2018):

1. Segmentação da mama: mapeamento da região da mama em relação ao fundo da imagem, bem como a localização da porção que compreende o músculo peitoral para imagens Mediolateral Oblíqua (MLO);
2. Análise da estrutura mamária: aplicação de um banco de filtros de Gabor na região em que se encontra a mama para enfatizar as fibras do tecido mamário;
3. Rastreamento das regiões suspeitas: levantamento das regiões que possuem fibras convergindo para um mesmo centro focal. Essas regiões, devido às características das DAM, irão compor a lista de regiões suspeitas;
4. Redução de falsos-positivos: como a quantidade de regiões suspeitas é muito grande, esta etapa busca eliminar regiões que não atingem um critério pré-estabelecido;
5. Extração de características: para as regiões remanescentes, são aplicados descritores de textura de modo a extrair características que possam discriminar regiões com DAM dentre aquelas normais;
6. Classificação: o classificador após ser treinado, irá selecionar aquelas regiões que são positivas para DAM.

Apesar das etapas conceitualmente bem definidas, cada uma apresenta suas limitações quando o objetivo é detectar regiões com DAM. Devido a grande heterogeneidade da estrutura mamária, a sua segmentação (detecção do músculo peitoral) bem como a análise da estrutura interna tem sido um grande desafio nos últimos anos (FERRARI et al., 2004; SAIDIN et al., 2012; CASTI et al., 2013; MUSTRA; GRGIC; RANGAYYAN, 2016). Após a segmentação, encontrar as áreas que correspondem a uma DAM, cuidando para que não ocorra o aumento de falsos-positivos, também tem sido foco de pesquisas (AYRES; RANGAYYAN, 2007; RANGAYYAN et al., 2008; BISWAS; MUKHERJEE, 2011; LIU et al., 2016a; MORRA; DELSANTO; GRAZZINI, 2018). Em uma segunda tentativa de eliminar regiões suspeitas, que não correspondem a DAM, o uso dos mais diversos extratores de características texturais tem sido explorado (AYRES; RANGAYYAN, 2003; RANGAYYAN; AYRES, 2006; BOZEK; DELAC; GRGIC, 2008; TANG et al., 2009; MINAVATHI; MURALI; DINESH, 2011; MOHAMMADI et al., 2013; LAKSHMANAN et al., 2014; LAKSHMANAN et al., 2015; KAMRA et al., 2016; LIU et al., 2016b; ZYOUT; TOGNERI, 2018).

A implantação da DBT na prática clínica possibilitou uma melhor detecção da DAM se comparada com a mamografia digital (PARTYKA; LOURENCO; MAINIERO, 2014; DIBBLE et al., 2018). Porém, vários casos de DAM ainda são perdidos dada a sutileza e complexidade dessa anomalia (SULEIMAN et al., 2016). Além disso, poucos apresentam propostas para a detecção da DAM em imagens de DBT (PALMA; BLOCH; MULLER, 2014; MORRA; DELSANTO; GRAZZINI, 2018).

Isso ocorre principalmente devido a quantidade de imagens que são geradas em um exame de DBT, geralmente mais de 30 imagens por incidência por mama. Uma opção seria aplicar o mesmo procedimento da mamografia 2D para cada imagem, o que acaba por não explorar a tridimensionalidade que a DBT oferece, além de aumentar consideravelmente o tempo de processamento e a quantidade de falsos-positivos gerados (MORRA; COCCIA; CERQUITELLI, 2018). Desse modo, a detecção de DAM em imagens de DBT ainda é um problema a ser resolvido (SULEIMAN et al., 2016; MORRA; DELSANTO; GRAZZINI, 2018). Como pôde ser observado, a detecção da DAM ainda é um problema em aberto que precisa ser investigado com maior atenção, pois há uma urgência de diagnósticos precisos justificados pelo alto nível de periculosidade associado a essa lesão.

Mesmo com o uso de exames avançados como a DBT, a taxa de detecção da DAM ainda está longe de ser satisfatória. Isso leva a hipótese de que, a união de um exame avançado com modelos computacionais mais robustos pode resultar em um avanço significativo para a detecção desse achado.

1.1 Objetivos

Após uma breve visão geral do estado da arte para a detecção de DAM em imagens mamográficas, é fato a necessidade de sistemas mais robustos para auxiliar os médicos em seus diagnósticos. Nesse cenário, o objetivo do presente trabalho é propor novas abordagens para a detecção automática da distorção arquitetural mamária (DAM), tanto em imagens de mamografia digital (mamografia 2D) como também para tomossíntese digital da mama (mamografia 3D). De maneira sucinta, os objetivos específicos dessa pesquisa são:

1. Melhorar a discriminação de regiões com DAM, dentre outros achados além do tecido normal, utilizando descritores de textura para micro-padrões locais;
2. Eliminar falsos-positivos em um esquema completo de detecção automática de DAM (*pipeline*);
3. Detectar a DAM de forma automática em imagens de tomossíntese digital, explorando a tridimensionalidade do exame;
4. Utilizar arquiteturas de aprendizagem profunda para a discriminação e reconhecimento de áreas com DAM.

Como resultado, espera-se que as técnicas aqui propostas possam contribuir para a detecção da distorção arquitetural mamária em imagens de mamografia digital bem como a tomossíntese digital da mama. Com um melhor desempenho na detecção dessa anomalia, os radiologistas poderão melhorar o diagnóstico precoce do câncer de mama, beneficiando significativamente os programas de rastreamento da doença e, conseqüentemente, a população em geral.

1.2 Organização da Tese

Após contextualizar o problema do câncer de mama e da detecção da DAM na mamografia, esta tese se divide nos seguintes capítulos:

Capítulo 2: são apresentados os procedimentos empregados no rastreamento do câncer de mama bem como os dois principais exames utilizados: a mamografia digital e a tomossíntese digital mamária, seções 2.1 e 2.2, respectivamente;

Capítulo 3: inicialmente será descrita a sequência de procedimentos (*pipeline*) que é usualmente utilizada para a detecção da DAM em imagens mamográficas. Cada uma das etapas apresentadas é detalhadamente explicada com ênfase nos pontos positivos e negativos. Por fim, serão apresentados os trabalhos que se correlacionam com este;

Capítulo 4: são descritos os bancos de imagens que contêm exames clínicos de mamografia 2D e 3D, montados em parceria com o InRAD (Instituto de Radiologia do Hospital das Clínicas da USP) e com a Clínica Eco & Mama;

Capítulo 5: neste são apresentadas todas as abordagens propostas, que em suma, visam superar as limitações apontadas no Capítulo 3;

Capítulo 6: são apresentadas as conclusões obtidas e o que se propõe para os trabalhos futuros;

Capítulo 7: são reproduzidos os artigos publicados durante o período de desenvolvimento desta tese.

O Câncer de Mama e o Rastreamento

Dentre as doenças que podem causar a morte principalmente da população feminina, o câncer de mama é a principal delas. Com o surgimento de 1,67 milhões de novos casos a cada ano apenas no Brasil, esta é uma doença que requer atenção especial dos governos e da população, para a sua detecção em estágio inicial e o tratamento necessário (WHO, 2019; INCA, 2019).

Os dados do Instituto Nacional do Câncer (INCA) corroboram os da OMS quanto a taxa de mortalidade das mulheres. Apenas no ano de 2018, estimou-se o surgimento de aproximadamente 60 mil novos casos no Brasil, dos quais aproximadamente 25% resultaram na morte da paciente (INCA, 2019).

A detecção do câncer de mama em estágio inicial e a aplicação do tratamento apropriado pode aumentar em até 30% a taxa de cura da paciente. Se a doença é detectada em estágio mais avançado de desenvolvimento, a chance de cura passa a ser cada vez mais remota (ELMORE et al., 2003; MAIDMENT, 2003; VERONESI et al., 2005). Com esse objetivo, os governos de diversos países desenvolvem programas de rastreamento e conscientização, para que as mulheres dentro da faixa etária de risco (de 40 a 50 anos, no Brasil) possam entender a seriedade da doença além de serem capazes de identificar sinais suspeitos (INCA, 2019).

Nos programas de rastreamento a paciente é incentivada a fazer a mamografia caso esteja dentro do grupo de risco que, além da idade, envolve o histórico dos casos de câncer na família. Quando o exame é feito, as imagens são analisadas por um médico radiologista, onde esse busca por sinais que possam indicar um câncer em desenvolvimento.

Apesar de haver tecnologias muito mais avançadas para o rastreio da doença, como a tomossíntese digital da mama (*Digital Breast Tomosynthesis* - DBT) que traz um maior benefício ao paciente, a mamografia 2D (analógica ou digital) ainda é o exame mais utilizado em todo o mundo (KARELLAS; VEDANTHAM, 2008; GLYNN et al., 2011).

2.1 Mamografia Digital

A mamografia digital de campo total (*Full Field Digital Mammography* - FFDM) é o exame mais indicado para o rastreio do câncer de mama na população feminina assintomática dentro da faixa de risco (a partir de 40 anos no Brasil) (KARELLAS; VEDANTHAM, 2008; GLYNN et al., 2011).

O exame de mamografia, ilustrado na Figura 2, consiste de uma radiografia da mama em que os fótons de luz irradiam do tubo de raios-X em direção ao detector e, desse modo atravessando a glândula mamária. Ao atingir o detector, os fótons de raio-X são quantizados formando uma imagem sem qualquer processamento de realce (formato *raw*). Um pós-processamento é aplicado na imagem *raw* para realçar a estrutura mamária facilitando a sua visualização e o diagnóstico (RUSSO, 2018). Além disso, essa mesma imagem processada é utilizada para a aplicação de sistemas computacionais para detecção de lesões.

Figura 2: Ilustração de exame mamográfico convencional, a Mamografia Digital de Campo Total (FFDM).

Fonte: Adaptado da *Wikipédia* - <https://en.wikipedia.org/wiki/Mammography>.

Apesar do seu intenso uso no rastreamento do câncer de mama em todo o mundo, a mamografia 2D apresenta uma grande limitação: a sobreposição de tecidos e estruturas mamárias na imagem. Isso ocorre devido à projeção de um volume tridimensional, nesse caso a mama, em um plano bidimensional, a imagem mamográfica (SECHOPOULOS, 2013).

Como ilustrado na Figura 3, o tubo de raios-X fica no lado oposto ao detector e, portanto, qualquer estrutura que esteja mais próxima do tubo irá sobrepor outras estruturas que estejam abaixo desta. O resultado, apresentado na Figura 4, é a oclusão total da “Estrutura 2” devido a “Estrutura 1” estar acima dela.

Figura 3: Ilustração da ocorrência da sobreposição de tecidos em mamografia 2D.

Fonte: O autor, 2019.

A sobreposição de tecidos é um problema crítico, sendo classificado como o principal responsável por erros no diagnóstico do câncer de mama pelos radiologistas (GLYNN et al., 2011; PARTYKA; LOURENCO; MAINIERO, 2014; PUJARA; HUI; WANG, 2019). Em casos reais, tal mascaramento faz com que muitos tumores fiquem encobertos por tecidos fibroglandulares e, por consequência, não sejam detectados no exame (BAKER; LO, 2011).

Figura 4: Ocorrência da sobreposição de tecidos em uma mamografia hipotética.

Fonte: O autor, 2019.

Por outro lado, a sobreposição de tecidos também pode “imitar” um padrão que leve o especialista a pensar que existe algo suspeito naquela região. Em ambas situações o

paciente é afetado de maneira negativa, seja por ter um problema que não foi detectado, seja por não haver qualquer doença e obter um exame que indique o contrário.

2.2 Tomossíntese Digital Mamária

A maior limitação da mamografia foi superada em 2011 com o início do uso clínico da Tomossíntese Digital da Mama (*Digital Breast Tomosynthesis* - DBT), também conhecida como “mamografia 3D” (VEDANTHAM et al., 2015).

No exame de DBT, ao contrário do que ocorre na mamografia digital, o tubo de raios-X se move descrevendo um arco de 15° ao redor da mama, como apresentado na Figura 5. Durante o movimento do tubo são adquiridas 15 imagens radiográficas (projeções). Essa configuração é característica do equipamento *Selenia Dimensions*® fabricado pela empresa americana Hologic, podendo vir a ser diferente de acordo com o modelo e/ou fabricante (SECHOPOULOS, 2013; PHILPOTTS; HOOLEY, 2018).

Figura 5: Ilustração de exame de tomossíntese digital mamária (DBT).

Fonte: O autor, 2019.

As projeções são processadas por *software* computacional de reconstrução tomográfica de modo a reconstruir a mama em sua tridimensionalidade. Esse volume é seccionado, representado pelas linhas acima da mama na Figura 5, e um novo conjunto de imagens (fatias) é produzido. Posteriormente, para que seja efetuado o laudo, apenas as imagens das fatias são apresentadas ao radiologista (VEDANTHAM et al., 2015; BAKER; LO, 2011; SECHOPOULOS, 2013).

Como cada secção corresponde a fatia da mama em uma determinada profundidade, facilmente o problema da sobreposição de tecidos é solucionado, como ilustra a Figura 6.

Estudos mostraram que a DBT além de minimizar o problema da sobreposição de tecidos mamários, possui maior sensibilidade e especificidade do que a mamografia convencional 2D. Além disso, espera-se que a DBT possa substituir a mamografia 2D como ferramenta de rastreamento do câncer de mama em um futuro próximo (RAFFERTY et al., 2013; RAY et al., 2015; DURAND et al., 2015)

Figura 6: Sequência de fatias mamográficas produzidas em um exame de DBT, no qual o problema da sobreposição de tecidos é solucionado.

Fonte: O autor, 2019.

2.3 Sistemas de Auxílio ao Diagnóstico

Com o intuito de auxiliar os radiologistas no diagnóstico precoce do câncer de mama, sistemas computacionais de auxílio ao diagnóstico (*Computer Aided Detection - CAD*) são normalmente utilizados na prática clínica (DOI, 2007; MAYO; LEUNG, 2018).

Os CADs, após uma análise da imagem, indicam de maneira automática regiões consideradas suspeitas na mamografia, desse modo fornecendo uma “segunda opinião” ao médico (DEAN; ILVENTO, 2006; YANG et al., 2007). Na Figura 7, é apresentada uma imagem de mamografia após ser analisada por um desses sistemas, na qual são apresentadas marcações que correspondem às regiões suspeitas encontradas.

Alguns estudos têm mostrado que os radiologistas, quando assistidos por sistemas CAD, são capazes de detectar na mamografia digital aproximadamente 90% das microcalcificações e nódulos. No entanto, quando se trata da detecção da DAM o desempenho dos sistemas CAD é insatisfatório e, portanto, muitos casos ainda são perdidos (BAKER et al., 2003; RANGAYYAN; BANIK; DESAUTELS, 2010; SULEIMAN et al., 2016).

Figura 7: Mamografica com marcações referente às regiões suspeitas encontradas pelo CAD ImageChecker® v.10.0.

Fonte: ImageChecker® *User Guide* (HOLOGIC, 2014).

De acordo com a problemática que envolve a detecção da DAM, tem-se estudado nos últimos anos novas alternativas para auxiliar na sua efetiva localização em estágio inicial. Nesse caso, uma das abordagens mais utilizadas atualmente são os sistemas computacionais baseados na análise dos gradientes orientados e descritores de textura (JASIONOWSKA; PRZELASKOWSKI, 2010; BANIK; RANGAYYAN; DESAUTELS, 2011; HANDA et al., 2012; BANIK; RANGAYYAN; DESAUTELS, 2013; RANGAYYAN et al., 2013; LAKSHMANAN et al., 2015; NETPRASAT; AUEPHANWIRIYAKUL; THEERA-UMPON, 2014; KAMRA et al., 2016; ZYOUT; TOGNERI, 2015).

Apesar da robustez das propostas mencionadas anteriormente, em sua grande maioria as bases de dados utilizadas são de imagens mamográficas 2D, algumas ainda utilizando

imagens de filme mamográfico digitalizado. Essas imagens possuem uma qualidade muito inferior se comparadas às tecnologias mais atuais como as imagens de mamografia digital (FFDM) e a tomossíntese digital mamária (DBT) (RANGAYYAN; AYRES, 2006; RANGAYYAN; BANIK; DESAUTELS, 2010; MENCATTINI et al., 2010; NETPRASAT; AUEPHANWIRIYAKUL; THEERA-UMPON, 2014; KAMRA et al., 2016).

Detecção Automática da Distorção Arquitetural na Mamografia

Para a detecção de lesões suspeitas na mamografia e, não obstante a distorção arquitetural, é necessário realizar uma sequência de etapas antes do sistema dar uma resposta. Essas etapas, muito similares aos *pipelines* clássicos de visão computacional, visam obter uma melhor acurácia na detecção, bem como evitar redundância de processamento.

O *pipeline* apresentado na Figura 8 reflete o que tem sido adotado na maioria dos trabalhos encontrados na literatura para a detecção automática da DAM em imagens de mamografia, sejam estas digitais ou não (RANGAYYAN; AYRES, 2006; BOZEK; DELAC; GRGIC, 2008; TANG et al., 2009; JASIONOWSKA; PRZELASKOWSKI, 2010; BANIK; RANGAYYAN; DESAUTELS, 2011; BISWAS; MUKHERJEE, 2011; ANAND; RATHANA, 2013; LAKSHMANAN et al., 2014; NETPRASAT; AUEPHANWIRIYAKUL; THEERA-UMPON, 2014; KAMRA et al., 2016; ZYOUT; TOGNERI, 2015; LIU et al., 2016a; LIU et al., 2016b; ZYOUT; TOGNERI, 2018).

Figura 8: Sequência de passos usualmente aplicados para a detecção da distorção arquitetural mamária.

Fonte: O autor, 2018.

A seguir, cada etapa da Figura 8 é brevemente apresentada e, para cada uma é apontada a seção na qual é melhor explicada:

1. **Segmentação da mama** (Seção 3.1): tem por objetivo extrair da imagem mamográfica apenas a região que compreende a mama, eliminando o fundo que não contém qualquer informação útil para auxiliar na detecção da DAM;
2. **Análise da estrutura mamária** (Seção 3.2): com a região da mama já definida pela etapa anterior, esse passo visa a aplicação de uma sequência de filtros para o realce da distribuição do tecido mamário;
3. **Rastreamento de regiões suspeitas** (Seção 3.3): regiões que possuem uma maior probabilidade de conter a DAM são aquelas nas quais os tecidos mamários convergem para um centro focal, a detecção das regiões que apresentam essa característica é executada nesta etapa;
4. **Redução de falsos-positivos** (Seção 3.4): devido a distribuição do tecido mamário, é natural que a etapa anterior encontre várias regiões suspeitas, esta etapa do *pipeline* tem a finalidade de mitigar o máximo possível das regiões que nada representam em termos de DAM;
5. **Extração de características** (Seção 3.5): apesar da etapa anterior ter reduzido uma boa parte das regiões suspeitas, algumas ainda persistem. Sendo assim, essas regiões precisam ser processadas de modo a discriminar efetivamente a possível região com DAM. Nessa etapa são aplicados descritores de textura em cada região suspeita, pois esses são os que podem melhor descrever a DAM.
6. **Classificação** (Seção 3.6): nesta etapa final, um classificador treinado com os vetores produzidos pelos descritores deverá ser capaz de descartar os falsos-positivos remanescentes.

Como já foi dito anteriormente, as seções que se seguem visam explorar com maior nível de detalhamento as etapas do *pipeline* da Figura 8. Além disso, na Seção 3.7, são apresentados os trabalhos que, de algum modo, se correlacionam com esta Tese de Doutorado.

3.1 Segmentação da Mama

As imagens de mamografia independente da incidência (CC ou MLO) possuem uma grande área de fundo, compreendida pela região escura das imagens apresentadas na Figura 9. Normalmente nas imagens disponíveis para o diagnóstico médico (formato processado) os pixels de fundo possuem valor 0. No entanto, longe de ser uma tarefa trivial, alguns problemas podem ser apontados (MUSTRA; GRGIC; RANGAYYAN, 2016):

1. A borda da mama não aparece bem definida na imagem;

2. A existência de *tags* de identificação no meio da área escura. Nas imagens da Figura 9, é possível visualizar essas *tags*;
3. Ruído e artefatos na imagem.

Figura 9: Exemplo de mamografias nas incidências (a) Craniocaudal (CC) e (b) Mediola-teral Oblíqua (MLO).

Fonte: O autor, 2019.

Quando se busca realizar a segmentação de alguma área em qualquer que seja a imagem, se esta apresenta um conjunto de estruturas que podem representar uma textura com alguma orientação, normalmente são empregados os filtros de Gabor (JAIN; FARROKHNI, 1991).

Os chamados *filtros de Gabor* ou *Wavelets de Gabor* são um conjunto de filtros simétricos e orientados, compostos por senoides modeladas por um “envelope” gaussiano (RANGAYYAN; AYRES, 2006). Um exemplo desses filtros é apresentado na Figura 10.

É possível calcular a orientação θ de uma textura qualquer em cada pixel (x, y) fazendo a convolução do *kernel* na imagem. Além de possuir uma orientação, o mesmo *kernel* também pode ter variação na escala, permitindo assim uma análise multi-resolução da estrutura de interesse. Considerando $\theta = -\pi/2$, o *kernel* do filtro de Gabor é definido

como:

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right] \cos(2\pi fx), \quad (1)$$

na qual, σ_x σ_y correspondem aos desvios-padrões nas direções x e y e, f é a frequência espacial.

Figura 10: Filtro de Gabor em vista (a) lateral e (b) superior.

Fonte: O autor, 2019.

Existem vários métodos para realizar a segmentação com diversas abordagens e processos de filtragem como por exemplo: segmentação baseada em *thresholding*, morfologia matemática, crescimentos de regiões, contornos ativos, dentre outros. Uma revisão de tais métodos pode ser lida na referência (MUSTRA; GRGIC; RANGAYYAN, 2016).

Dentre os trabalhos apontados anteriormente, o proposto em Casti *et al.* (2013) é o único que fez uso de imagens de mamografia digital. O método apresentado utiliza o filtro de Gabor multidirecional juntamente com o método de segmentação de Otsu (GONZALEZ; WOODS, 2008) para obter uma estimativa inicial das limítrofes da mama. Esta estimativa inicial é refinada com o uso do cálculo da distância Euclidiana entre as duas imagens iniciais, seguido por uma nova etapa de binarização com *threshold* local e, finalizada com a aplicação de um método para supressão de falsas bordas que possam aparecer. Os resultados reportados, em termos de completude e exatidão ficaram em torno de 99%.

O fundo é uma região que está presente nas imagens de ambas incidências, CC ou MLO. No entanto, para imagens MLO outra região precisa ser detectada e removida, a área do músculo peitoral. Esta região é compreendida pela parte mais clara no canto

superior direito da imagem apresentada na Figura 9b, não faz parte da glândula mamária e é composto por outro tipo de tecido.

Foco de várias outras pesquisas, a segmentação do músculo peitoral tem sido um desafio (FERRARI et al., 2004; SAIDIN et al., 2012; CASTI et al., 2013; MUSTRA; GRGIC; RANGAYYAN, 2016). O trabalho apresentado em (FERRARI et al., 2004), é um dos mais utilizados na literatura. Nele é proposto um conjunto de métodos que utiliza como base a resposta do filtro de Gabor em várias direções e escalas. A magnitude juntamente com os componentes de fase são combinados e pós-processados para formação de uma borda completa, sem buracos, ou o descarte no caso de serem pontos desconexos.

Com a aplicação dos métodos propostos em (CASTI et al., 2013) e (FERRARI et al., 2004) é possível realizar a segmentação das imagens nas duas incidências. As linhas em verde, presente nas imagens da Figura 11, mostram os resultados após a aplicação dos métodos apresentados anteriormente.

Figura 11: Exemplo de mamografias nas incidências (a) Craniocaudal (CC) e (b) Mediolateral Oblíqua (MLO), ambas segmentadas nas regiões que compreende a mama.

Fonte: O autor, 2019.

Desse modo, após a aplicação dos procedimentos referentes à segmentação nas mamografias apresentadas nas figuras 12a e 12d, é possível obter as máscaras binária das figuras 12b e 12e. As imagens estarão prontas para o processamento quando for

efetuada a multiplicação da máscara binária com a imagem original, resultando na imagem segmentada, como mostram as figuras 12c e 12f.

Figura 12: Mamografias nas incidências CC e MLO, (a-d) sem qualquer processamento, (b-e) a máscara binária gerada e (c-f) a imagem após a finalização do processo de segmentação da mama.

Fonte: O autor, 2019.

3.2 Análise da Estrutura Mamária

Com a mama segmentada, o próximo passo é o realce das estruturas internas. Uma avaliação de diversos operadores para a detecção da estrutura mamária, como o operador de linha, filtros de Gabor e os filtros direcionais (da classe *steerable*), mostrou que o filtro de Gabor obteve os melhores resultados (AYRES; RANGAYYAN, 2007).

O método proposto por Ayres e Rangayyan (2007), posteriormente melhorado por Casti *et al.* (2016), é composto da aplicação de 18 filtros direcionais obtidos após a rotação do *template* gerado pela Equação 1, apresentada na Seção 3.1. O processamento da mamografia com cada filtro direcional gera uma imagem diferente, a recuperação do valor de maior magnitude de cada pixel (x, y) consiste na observação da mesma posição nas outras imagens. Com o objetivo de remover eventuais ruídos e detalhes finos que possam atrapalhar as etapas futuras, é recomendado um pós-processamento com a aplicação de um filtro Gaussiano (RANGAYYAN; AYRES, 2006; RANGAYYAN; BANIK; DESAUTELS, 2010).

Considerando a mamografia segmentada da Figura 12f, após a aplicação da abordagem apresentada nesta seção, o resultado pode ser observado na Figura 13. Na imagem resultante, as mais diversas estruturas da mama foram realçadas para que, posteriormente, se encontrem aquelas que possam vir a ser uma DAM.

Figura 13: Resultado do realce da estrutura mamária da imagem em 12f.

Fonte: O autor, 2019.

3.3 Rastreamento de Regiões Suspeitas

Regiões suspeitas de serem DAM são aquelas onde as fibras do tecido mamário tendem a convergir para uma região central. A detecção desse tipo de padrão tem como

ponto de partida o resultado da etapa anterior, ou seja, a imagem cujas estruturas da mama foram realçadas.

Os valores de intensidade de uma imagem podem ser caracterizados por meio de um *Monge patch*, $I : D \subset \mathcal{R}^2 \rightarrow \mathcal{R}^3$, definido como (AYRES; RANGAYYAN, 2003):

$$I(x, y) = [x, y, h(x, y)],$$

sendo D o domínio bidimensional (x, y) e $h(x, y) : D \rightarrow [0, 1]$ é uma função que, para cada coordenada espacial (x, y) associa um valor de intensidade $h \in [0, 1]$ (ABBENA; SALAMON; GRAY, 2006). Como h possui suas derivadas parciais dentro do intervalo $[0, 1]$ e a sua superfície é suave em qualquer ponto, devido a filtragem gaussiana, então uma nova variável K pode ser definida como (AYRES; RANGAYYAN, 2003; CASTI et al., 2016):

$$K = \det [\text{Hess}(h)],$$

no qual $\text{Hess}(h)$ é o operador Hessiano aplicado em h e, K é uma curvatura gaussiana podendo ter a sua topologia analisada de acordo com os valores encontrados em cada ponto. Seja $\mathbf{p} = (x_0, y_0)$ um ponto qualquer em h então, a sua topologia pode ser definida de acordo com o valor de K correspondente:

$$\begin{aligned} K(x_0, y_0) &> 0, & \text{então } \mathbf{p} \text{ é elíptico,} \\ K(x_0, y_0) &< 0, & \text{então } \mathbf{p} \text{ é hiperbólico,} \\ K(x_0, y_0) &= 0, & \text{então } \mathbf{p} \text{ é parabólico ou planar.} \end{aligned}$$

Se uma base ortonormal de autovalores para $\text{Hess}(h)$ é escolhida, então a curvatura gaussiana K pode ser representada por meio do produto desses autovalores, λ_1 e λ_2 . Assumindo que uma coordenada $(\tilde{x}, \tilde{y})$ satisfaz a condição $\lambda_{1,2}(\tilde{x}, \tilde{y}) < 0$, então ela será considerada como parte de uma região onde as estruturas da mama convergem (os vetores gradientes naquela região apontam para um mesmo centro focal) e, portanto, é uma região suspeita de DAM.

O resultado do processamento descrito nesta seção pode ser visualizado na imagem da Figura 14. O processo de segmentação da mama resultou na área delimitada em verde e, em seguida, a detecção de regiões de convexidade foi executado. Tais regiões, que passam a ser consideradas suspeitas estão destacadas em vermelho, nas quais cada centroide é marcado por um asterisco branco (RANGAYYAN; BANIK; DESAUTELS, 2010; RANGAYYAN et al., 2013).

Figura 14: Resultado da imagem apresentada na Figura 12d, após a segmentação e rastreamento de regiões suspeitas.

Fonte: O autor, 2019.

3.4 Redução de Falsos-Positivos

A etapa anterior produz uma grande quantidade de regiões suspeitas que, considerando uma mamografia com apenas uma região com DAM, resulta em uma enorme quantidade de falsos-positivos: regiões inicialmente consideradas suspeitas mas que não são a lesão. Devido a isso, executa-se algum procedimento para a redução automática dessas regiões. Estratégias para esse fim têm sido propostas ao longo dos anos, no entanto, este ainda é um problema em aberto (RANGAYYAN; BANIK; DESAUTELS, 2010; RANGAYYAN et al., 2013; OLIVEIRA et al., 2018; OLIVEIRA et al., 2019).

A Figura 15 ilustra os procedimentos que usualmente são aplicados (RANGAYYAN; BANIK; DESAUTELS, 2010; RANGAYYAN et al., 2013; CASTI et al., 2016). Considerando duas imagens distintas, que já passaram por todo processamento até o levantamento das regiões suspeitas, essas regiões agora são ordenadas (em ordem decrescente) de acordo com a sua intensidade média. Na sequência, com um *threshold* fixo pré-definido, a lista com as regiões é reduzida de modo a descartar aquelas que possuem intensidade abaixo do valor de corte.

Um primeiro problema que surge com essa abordagem é o descarte da região com DAM (verdadeiro-positivo) juntamente com os falsos-positivos, como mostra a área

sombreada referente a Imagem #2, na Figura 15. Esse tipo de situação compromete todo o resto do *pipeline*, pois a região com DAM foi perdida.

Figura 15: Exemplo da ordenação e corte da lista de regiões suspeitas.

Fonte: O autor, 2019.

Por outro lado, é possível especificar um valor de *threshold* alto como uma tentativa de evitar a perda da região com DAM. Isso leva ao segundo problema que vai contra ao objetivo dessa etapa: remoção de falsos-positivos. Assim, uma grande quantidade de regiões ainda são mantidas e, portanto, descartando a usabilidade dessa etapa.

Além dos problemas já citados, é natural que para cada imagem uma quantidade diferente de regiões suspeitas são encontradas. Desse modo, os problemas já citados podem, ou não, ocorrer devido a escolha errônea do valor de corte.

Mesmo que a aplicação do *threshold* tenha sido bem sucedida, ainda podem restar muitas regiões suspeitas. Uma vez que a DAM não tem borda definida e seu contraste pode não ser muito relevante em relação à sua vizinhança, resta apenas descrevê-la em termos de sua textura (RANGAYYAN; BANIK; DESAUTELS, 2010; OLIVEIRA et al., 2017; LIU et al., 2016b; LIU et al., 2017).

3.5 Extração de Características

Em visão computacional, métodos “descritores” ou “extratores de características” são aqueles que ao serem aplicados em uma imagem ou região de interesse, geram um vetor numérico conhecido por “vetor de características”. Tais características podem ser histogramas dos tons de cinza, informações de textura, informações geométricas, dentre outras (WILSON; RITTER, 2000; DAVIES, 2012). Em suma, as informações extraídas são propriedades de componentes das imagens (objetos e/ou pessoas) capazes de discriminar o objeto de interesse dentre outros (WILSON; RITTER, 2000).

Para o caso da DAM, ela não possui borda bem definida e muitas vezes não há diferença significativa de tons de cinza entre ela e o tecido mamário normal. Dessa forma,

não se pode usar descritores de forma nem de cor para descrever uma DAM. Nesse caso, a classe de descritores mais utilizados é a dos descritores de textura.

Os métodos que são capazes de descrever uma textura são normalmente classificados em quatro categorias de acordo com o método de análise empregado, sendo eles: estatístico, geométrico, baseado em modelos e, baseado no processamento de sinais (DAVIES, 2012). Tradicionalmente, as abordagens estatísticas são as mais bem sucedidas, sendo as matrizes de coocorrência, proposta por (HARALICK, 1979) a mais famosa delas.

Apesar de outros métodos terem sido propostos, a grande maioria não consegue discriminar muito bem texturas em imagens naturais ou são muito complexos para aplicações de tempo real. A última grande contribuição veio com a proposta de um método chamado *Local Binary Pattern* (LBP), que apesar da sua baixa complexidade, consegue descrever texturas de maneira muito satisfatória (OJALA; PIETIKAINEN; MAENPAA, 2002; PIETIKÄINEN et al., 2011).

3.5.1 Descritores Baseados em Micro-padrões Local

O *Local Binary Pattern* (LBP) é um operador de textura de baixa complexidade que rotula, com um número binário, cada pixel analisado com base na diferença dos tons de cinza dos seus vizinhos. Sendo assim, o LBP pode ser visto como uma abordagem que emprega a divergência estatística e a análise estrutural da textura (OJALA; PIETIKÄINEN; HARWOOD, 1996).

Figura 16: Vizinhança (de g_1 até g_8) de um pixel arbitrário g_0 considerando raio unitário.

Fonte: O autor, 2018.

Em sua formulação original proposta em (OJALA; PIETIKÄINEN; HARWOOD, 1996), ao analisar o pixel g_0 o LBP considera os $P = 8$ pixels em sua vizinhança circular de raio 1, como é ilustrado na Figura 16.

Com base em g_0 a Equação 2 é definida da seguinte maneira:

$$LBP = \sum_{i=1}^P S(g_0, g_i) 2^{i-1}, \quad (2)$$

na qual g_0 é comparado com todos os seus P vizinhos por meio da função *Heavyside* apresentada na Equação 3:

$$S(g_0, g_i) = \begin{cases} 1, & \text{se } g_i \geq g_0 \\ 0, & \text{caso contrário.} \end{cases} \quad (3)$$

Como pode ser observado na Equação 2, ao considerar 8 vizinhos o termo 2^{i-1} garante um LBP de valores que variam de 0 até 255. Ao finalizar o cálculo do LBP para cada pixel, um histograma do padrão binário é montado e utilizado como descritor.

Para exemplificar considere a imagem da Figura 17a, que possui áreas com variados tipos de textura, bem como áreas uniformes. A aplicação do LBP é uma operação pontual, sendo assim, tem-se como resultado a imagem de códigos LBP ou mapa de textura apresentada na Figura 17b. A partir dessa imagem, o histograma da Figura 18 é gerado para utilização como vetor descritor.

Figura 17: Exemplo de aplicação do LBP na imagem (a), gerando (b) o mapa de códigos LBP.

Fonte: O autor, 2018.

Ao se comparar as imagens das figuras 17a e 17b, é possível notar que áreas uniformes possuem um código LBP alto, isso quer dizer que ao comparar o pixel central (g_0 na Equação 3) com os seus i vizinhos ambos possuem a mesma intensidade fazendo com que a aplicação da função *Heavyside* $S(\cdot)$ retorne 1.

Apesar da imagem 17b parecer binária, seus tons de cinza (códigos LBP) variam de 0 a 255, como comentado anteriormente. Essa situação é melhor ilustrada por meio da

Figura 18: Histograma de códigos LBP.

Fonte: O autor, 2018.

análise do seu respectivo histograma, Figura 18, no qual se observa um pico maior para intensidades próximas de 255 indicando que a imagem possui uma maior quantidade de áreas uniformes que são brancas.

O LBP vem sendo utilizado nas mais diversas aplicações com a adoção de outras configurações como por exemplo, considerando outros valores de raio e quantidade de vizinhos. Além disso, outras propostas buscam mudar o modo de como os pixels vizinhos serão analisados e, inclusive como a diferença entre os pixels é mapeada, como acontece com o *Local Mapped Pattern* (LMP) (BOAVENTURA; GONZAGA, 2009; VIEIRA et al., 2012; CHIERICI et al., 2013; FERRAZ et al., 2014; FERRAZ; PEREIRA; GONZAGA, 2014).

3.5.1.1 *Local Mapped Pattern*

O *Local Mapped Pattern* (LMP) é um descritor que, assim como o LBP, extrai informações relacionadas a textura em uma dada região da imagem. De fato, o LMP se trata de uma generalização do tradicional LBP (BOAVENTURA; GONZAGA, 2009; VIEIRA et al., 2012). A maior diferença entre os dois métodos está na função de mapeamento utilizada que, enquanto o LBP faz uso apenas de uma função degrau (*Heavyside*), o LMP pode ser utilizado com qualquer função de mapeamento. Tal recurso tem gerado resultados superiores na classificação de diversos tipos de textura, inclusive as coloridas, rotacionadas e com variação de iluminação (BOAVENTURA; GONZAGA, 2009; VIEIRA et al., 2012; CHIERICI et al., 2013; FERRAZ et al., 2014; FERRAZ; PEREIRA; GONZAGA, 2014; OLIVEIRA et al., 2016).

Com o objetivo de expandir os recursos do LBP, o LMP possibilita o uso de qualquer função de mapeamento e permite uma variação na quantização do resultado, fazendo com que o vetor descritor possa ser mapeado em um histograma de B bins (BOAVENTURA; GONZAGA, 2009; VIEIRA et al., 2012). Sendo assim, o LMP parte inicialmente da formulação do LBP (Equação 2) que foi reescrita da seguinte maneira:

$$LMP = round \left(\frac{\sum_{i=1}^W f_g \cdot P(i)}{\sum_{i=1}^W P(i)} (B - 1) \right), \quad (4)$$

na qual os seguintes parâmetros são estabelecidos:

W : quantidade de vizinhos;

f_g : função de mapeamento;

P : matriz de pesos, também podendo ser obtida pelo termo 2^{i-1} da Equação 2;

B : quantidade de bins no histograma descritor final.

Considerando uma janela 3×3 e, portanto raio unitário e $W = 8$, a matriz de pesos P , apresentada anteriormente como o termo 2^{i-1} da Equação 2, é explicitamente definida como mostra a Equação 5. Sendo assim, com uma matriz de pesos pré-definida, a magnitude das diferenças entre dois ou mais pixels pode ser alterada de modo a obter uma melhor resposta de acordo com o problema a ser resolvido:

$$P(i) = \begin{bmatrix} 1 & 2 & 4 \\ 128 & \mathbf{0} & 8 \\ 64 & 32 & 16 \end{bmatrix}, \quad (5)$$

na qual a disposição dos pixels continua a mesma apresentada na Figura 16, sendo $\mathbf{0}$ o pixel central.

Talvez o recurso mais interessante que o LMP apresenta em relação ao LBP é a possibilidade de mudança da função de mapeamento f_g , podendo esta ser alterada de acordo com o problema a ser abordado. Portanto, em sua formulação inicial, o LMP apresenta o uso da função *Sigmoid*, apresentada na Equação 6, ao invés da função *Heavyside* (Equação 3):

$$f_g = \frac{1}{1 + e^{\frac{-(g_i - g_0)}{\beta}}}, \quad (6)$$

na qual, β é um parâmetro que define a inclinação da curva *Sigmoid*, cujo o resultado da variação dos valores é observado no gráfico da Figura 19. Além disso, observa-se que no caso particular onde $\beta = 0$, a função *Sigmoid* tende a convergir para a função *Heavyside*.

Figura 19: Gráfico da função *Sigmoid* com diferentes parâmetros β .

Fonte: O autor, 2018.

Portanto, a configuração do parâmetro β permite o ajuste preciso do descritor tornando-o mais robusto especificamente para o problema a ser resolvido. Além disso, o número de *bins* do histograma descritor também pode ser otimizado variando-se o parâmetro B . Em adendo, o LMP permite também avaliar diferentes “funções de mapeamento” como por exemplo a função triangular que se mostrou eficiente para a detecção de bordas em imagens digitais (BOAVENTURA; GONZAGA, 2009).

Apesar dos diversos recursos que o LMP apresenta em relação a outros descritores, é necessário ressaltar que a configuração ótima desses parâmetros é uma tarefa que exige muito recurso computacional. Uma vez que se tenha um parâmetro alterado, seja o número de *bins* ou o β , para que esta configuração seja devidamente testada, a etapa de classificação deve ser executada.

3.6 Classificação

Talvez a última etapa da maioria dos *pipelines* de visão computacional é a “Classificação”. É nela que as diferentes classes de imagens serão separadas, após um efetivo treinamento do classificador utilizado, se este for treinável. Para o caso de detecção da DAM, esta abordagem “clássica” ainda é válida visto que, ao fim do processo (apresen-

tado inicialmente na Figura 8), espera-se que apenas uma região (destacada em amarelo) permaneça, pois ela corresponde a localização da distorção arquitetural.

Dentre os diversos classificadores existentes na literatura, as Redes Neurais Artificiais são de longe as mais utilizadas (DAVIES, 2012; FERNÁNDEZ-DELGADO et al., 2014). A Figura 20, ilustra um *Perceptron*, a forma mais simples de uma Rede Neural pois possui apenas um neurônio e trabalha como um classificador linear, podendo separar padrões em duas classes.

Figura 20: Neurônio artificial *Perceptron*.

Fonte: DAVIES, 2012.

De uma maneira simplificada, tomando como exemplo o *Perceptron* da Figura 20, a rede neural recebe como entrada um vetor de características com N elementos, $x = (x_1, x_2, x_3, x_4, \dots, x_N)$. A sua saída consiste na geração de um valor resultante da equação: $\sum_{i=1}^N w_i x_i$, após a aplicação de uma função de ativação f , neste caso a *Heaviside*. A saída linear do classificador é descrita na Equação 7:

$$d = \sum_{i=1}^N w_i x_i = \sum_{i=1}^N w_i x_i + w_0 - \theta, \quad (7)$$

e portanto, a sua resposta final é dada após a aplicação da função f em d , como mostra a Equação 8:

$$y = f(d) = f\left(\sum_{i=1}^N w_i x_i\right) \quad (8)$$

O conceito de *Perceptron* apresentado anteriormente é válido para o entendimento inicial de uma rede neural artificial. No entanto, a classificação de distorção arquitetural requer o uso de redes mais robustas, possuindo uma maior quantidade de neurônios organizados em diversas *camadas ocultas*, configurando assim um classificador não linear.

Vale ressaltar que, apesar da simplicidade da rede *Perceptron*, é possível notar que este pode ser expandido e modificado de acordo com a necessidade do problema a ser resolvido. Além da topologia lógica da rede (quantidade de neurônios e camadas)

existem diversas técnicas de treinamento que buscam sempre os melhores resultados, uma convergência mais rápida e consequentemente um menor custo computacional.

3.7 Trabalhos Correlatos

Nesta seção serão apresentados alguns dos trabalhos que mais se relacionam com este para a detecção da distorção arquitetural mamária. Sucintamente listados na Tabela 1, é possível notar que existe uma grande heterogeneidade dentre as diversas propostas apresentadas, o que ressalta a motivação de que a detecção da DAM ainda é um problema a ser resolvido.

Nemoto *et al.* (2009) propuseram um *pipeline* completo para a análise de 25 mamografias digitalizadas. Após uma etapa de pré-processamento, foram extraídas 9 características das regiões suspeitas: 6 delas (4 medidas de contraste, energia e *skewness*) obtidas na imagem de tons de cinza; 2 (*skewness* e o segundo momento) calculadas em uma imagem após a aplicação do filtro Laplaciano e; uma última característica foi a medida de distância entre a lesão e a pele da mama. Após uma classificação com a distância de Mahalanobis, o resultado obtido teve uma sensibilidade de 80%.

Utilizando imagens no domínio da frequência, Jasionowska *et al.* (2010) propuseram uma abordagem na qual 13 características foram extraídas de 292 imagens digitalizadas: 34 com DAM e 258 sem. Algumas dessas características (contraste, correlação energia, entropia e homogeneidade) são descritores de Haralick, outras (variância, desvio padrão, *kurtosis*, energia e entropia) são medidas estatísticas e, por fim, duas (inclinação e ponto de intersecção com o eixo-*y*) são baseadas na aproximação do espectro angular. Após uma classificação com *Support Vector Machine* (SVM), a sensibilidade reportada foi de 68%.

Minavathi *et al.* (2011) propuseram um método de distinção de ROIs no qual foram extraídas características referente a informação textural baseadas em vetores gradiente e na imagem de intensidades. Um total de 291 ROIs, 42 com DAM e 249 sem, das bases mini-MIAS e DDSM foram utilizados. O processo de classificação foi avaliado em dois classificadores, *perceptron* multi-camadas e SVM com *kernel* gaussiano. O melhor resultado obtido, utilizando o classificador SVM, foi com uma sensibilidade de 94,7%, especificidade de 89,6% e 0,90 de *Area Under the Curve* - Área Sob a Curva (AUC) *Receiver Operating Characteristic* (ROC).

Rangayyan *et al.* (2013) propuseram 13 características resultantes de medições referentes ao padrão espiculado que a DAM representa. Para isso foram utilizadas 158 mamografias digitalizadas, 106 com DAM e 52 sem. O resultado alcançado, com uma rede *perceptron* multi-camadas, obteve uma sensibilidade de 80% e AUC de 0,75.

Mohammadi *et al.* (2013) apresentaram o uso do descritor de textura *Monogenic*

Binary Coding (MBC) para discriminar regiões com DAM em imagens de mamografia digitalizada. O *dataset* utilizado foi o IRMA, uma versão baseada no DDSM, composto de 110 ROIs com DAM e 110 considerados normais. A performance reportada em termos de acurácia, sensibilidade e especificidade, foram de 91,2%, 87,5% e 95%, respectivamente.

Palma *et al.* (2014) propuseram um método para detecção de nódulos e DAM em imagens de DBT. A parte do *dataset* correspondente a DAM é composto de apenas 3 volumes. As características extraídas são relacionadas a análise da convergência de tecidos no que os autores chamaram de “abordagem *a contrario*”. Após uma etapa de classificação com SVM, os autores obtiveram uma sensibilidade de 80%.

Um método baseado no filtro SUSAN foi proposto por Lakshmanan *et al.* (2015) para a detecção de regiões suspeitas em 100 imagens de mamografia digital. A descrição das regiões com DAM fez uso de uma abordagem cujo objetivo é a contagem da excentricidade de estruturas que convergem para um único ponto central. Os autores reportaram uma sensibilidade de 94%.

Kamra *et al.* (2016) apresentaram uma proposta na qual 284 características de textura foram extraídas de mamografias. Três *datasets* de imagens digitalizadas foram utilizados (DDSM, mini-MIAS e uma base privada) totalizando 348 imagens (165/183 com/sem DAM). De cada ROI, 280 características foram extraídas com base nas matrizes de coocorrência, 2 são baseadas em fractais e outras 2 foram extraídas do espectro de potência de Fourier. Após uma etapa de redução de dimensionalidade, a classificação se deu com o SVM. Os resultados são reportados de acordo com o *dataset*, no entanto o DDSM teve o melhor resultado: acurácia e sensibilidade de 93,3%, especificidade de 92% e AUC ROC de 95%.

Casti *et al.* (2016) propuseram um *pipeline* completo para a classificação de anomalias entre benignas e malignas. A descrição contou com a extração de 7 características (área, compacidade, alongamento, distância, direcionalidade, direcionalidade ponderada e dispersão). Dentre outros resultados, aqueles referentes a classificação da DAM, utilizou um conjunto de 116 imagens todas com DAM (98 malignas e 24 benignas). Com o classificador *Fisher Linear Discriminant Analysis*, reportaram uma sensibilidade de 81%.

O foco do trabalho proposto por Liu *et al.* (2016b) é o classificador chamado *Twin Bounded SVM* (TBSVM) para separar um conjunto de 138 ROIs (69/69 com/sem DAM) da base DDSM. Para isso foram calculadas as características: 228 advindas das matrizes de coocorrência; 11 são medidas relacionadas a padrões espiculados e; por fim, as entropias de Renyi e Tsallis. A sensibilidade reportada neste caso foi de 90%.

Como pôde ser observado, algumas dessas propostas apresentam um *pipeline* completo, automático, para a detecção da DAM. Por outro lado, a grande maioria dos trabalhos propõem a classificação em nível de Regiões de Interesse (ROI) que foram

Tabela 1: Trabalhos mais relevantes que buscam a detecção da DAM. Legenda: FPs - Falsos Positivos; N.D. - Não Disponível. Os números no formato “X/Y” correspondem as quantidades “com/sem” DAM.

Autor(es)	Ano	Dataset	FPs	Sensibilidade	AUC	Acurácia
Nemoto <i>et al.</i>	2009	25 mamografias digitalizadas com DAM - 25/0	0,80	80%	N.D.	N.D.
Jasionowska <i>et al.</i>	2010	292 ROIs do DDSM - 34/258	0,86	68%	N.D.	N.D.
Minavathi <i>et al.</i>	2011	291 ROIs do DDSM e mini-MIAS - 42/249	N.D.	94%	0,90	N.D.
Rangayyan <i>et al.</i>	2013	158 mamografias digitalizadas - 106/52	3,40	80%	0,75	N.D.
Mohammadi <i>et al.</i>	2013	220 ROIs do IRMA (DDSM) - 110/110	N.D.	87%	N.D.	91%
Palma <i>et al.</i>	2014	101 conjuntos de DBT - 3/48	1,31	80%	N.D.	N.D.
Lakshmanan <i>et al.</i>	2015	100 mamografias digitais - 100/0	N.D.	94%	N.D.	N.D.
Kamra <i>et al.</i>	2016	348 ROIs do DDSM e mini-MIAS - 165/183	N.D.	93%	0,95	92, 9%
Casti <i>et al.</i>	2016	116 mamografias do DDSM, todas com DAM - 98 malignas e 24 benignas	7,65	80%	N.D.	N.D.
Liu <i>et al.</i>	2017	138 ROIs do DDSM - 69/69	6,70	90%	N.D.	N.D.

previamente anotadas por radiologistas. A grande maioria, utiliza imagens de mamografia digitalizada dos *datasets* públicos mini-MIAS (SUCKLING et al., 1994) e/ou DDSM (HEATH et al., 2000) que, se comparada a mamografia digital, possuem uma qualidade muito inferior. Além disso, pouquíssimos são os estudos que utilizam imagens de DBT (MORRA; DELSANTO; GRAZZINI, 2018).

Materiais e Métodos

Neste capítulo serão apresentados os bancos de imagens utilizados para o desenvolvimento deste trabalho, assim como as soluções propostas para melhorar o desempenho dos sistemas de detecção automática de DAM na mamografia digital e também na tomossíntese mamária.

4.1 Bancos de Imagens Mamográficas Clínicas

O desenvolvimento deste trabalho requer o uso de imagens clínicas de pacientes reais que tenham sido diagnosticadas com a distorção arquitetural mamária. Como o acesso a imagens clínicas é restrito, por serem dados confidenciais, ao longo do desenvolvimento do projeto vários bancos de imagens, com uma quantidade variada de imagens, foram utilizados. Inicialmente foi utilizado um banco de imagens de mamografias digitalizadas publicamente disponível na internet, o DDSM (HEATH et al., 2000).

Após algum tempo, foi firmada parceria com o Instituto de Radiologia da Faculdade de Medicina da USP (InRAD) e a Clínica Eco & Mama, que possibilitaram o acesso às imagens digitais tanto de mamografia de campo total (FFDM) quanto de tomossíntese mamária (DBT). Os dois bancos de imagens, digitalizadas e digitais, serão comentados nas seções 4.1.1 e 4.1.2, respectivamente.

4.1.1 *Digital Database for Screening Mammography*

O *Digital Database for Screening Mammography* (DDSM) é um banco de imagens contendo mamografias provenientes de equipamentos analógicos (tela-filme), que posteriormente foram digitalizadas e disponibilizadas *online* a partir do ano 2000 (HEATH et al., 2000). Esse banco foi montado pelas instituições americanas: *University of South Florida*, *Sandia National Laboratories* e *Washington University School of Medicine*, e é distribuído gratuitamente por meio da internet sem qualquer necessidade de registro.

Na Figura 21, são apresentadas algumas imagens que compõem o banco DDSM. São totalizados 2620 exames, de ambas as mamas nas incidências CC e MLO, contendo as anomalias: nódulo, microcalcificação, assimetria bilateral e distorção arquitetural. Cada exame, conjunto de imagens, é acompanhado de laudo médico bem como a anotação da região onde a lesão se concentra.

Figura 21: Imagens mamográficas do banco DDSM nas incidências (a) Mediolateral Obliqua (MLO) e (b) Craniocaudal (CC).

Fonte: Base de imagens DDSM (HEATH et al., 2000).

A grande quantidade de dados que estão agrupados no DDSM tem feito com que essa base venha sendo utilizada nos mais diversos trabalhos até os dias atuais (RANGAYYAN; BANIK; DESAUTELS, 2010; HANDA et al., 2012; NETPRASAT; AUEPHANWIRIYAKUL; THEERAMPON, 2014; KAMRA et al., 2016; ZYOUT; TOGNERI, 2015). Fatores como o intenso uso e o acesso gratuito motivaram a sua utilização nas etapas iniciais deste trabalho, uma vez que ainda não havia sido firmada parceria com hospital ou clínica para a aquisição de imagens digitais FFDM e DBT.

4.1.2 Imagens Digitais

Visando suprir a necessidade de imagens digitais e principalmente imagens de DBT, inicialmente foi firmada parceria com o Instituto de Radiologia da Faculdade de Medicina da USP (InRAD) e, posteriormente com a Clínica Eco & Mama situada na cidade de São Carlos/SP. Para ambos os centros de diagnóstico e imagem foram obtidas

as respectivas aprovações dos Comitês de Ética em Pesquisa (CEP) para o InRAD¹ e Clínica Eco & Mama².

Figura 22: Imagens clínicas de mamografia FFDM (2D) do banco de imagens do InRAD, nas incidências (a,c) Mediolateral Oblíqua (MLO) e (b,d) Craniocaudal (CC).

Fonte: O autor, 2018.

Por meio das parcerias, após a aprovação de comitê de ética em pesquisa, foi possível ter acesso e montar um banco de imagens (mamografia 2D e 3D) laudadas de

¹InRAD: CAAE #56699016.7.0000.0065)

²Clínica Eco & Mama: CAAE #80264717.3.0000.5393)

diversas anomalias, entre elas a DAM. Como ambos centros utilizam o mesmo modelo de mamógrafo, *Selenia Dimensions*® fabricado pela empresa americana Hologic, serão mostradas na Figura 22 apenas as imagens referentes ao banco de imagens do InRAD.

Figura 23: Imagens clínicas de DBT (fatias) utilizadas nesse trabalho, nas incidências (a) Mediolateral Oblíqua (MLO) e (b) Craniocaudal (CC). Para cada volume é apresentada uma fatia intermediária.

Para os casos nos quais o exame empregado foi a DBT, a quantidade de imagens é muito maior do que a de exames FFDM pois, para cada uma das 4 imagens de FFDM (duas mamas nas incidências CC e MLO) existe um volume de imagens associado. Cada um desses volumes é composto por uma quantidade de imagens (fatias) que variam de acordo com a espessura da mama comprimida durante o exame. Por exemplo, para uma mama com 5 cm de espessura, quando comprimida, serão geradas 50 fatias (uma para cada milímetro) após o processo de reconstrução. Na Figura 23, são apresentados dois volumes nos quais uma fatia intermediária é apresentada em tamanho maior.

As imagens coletadas nos centros de imagem, apesar de possuírem seus respectivos laudos, não possuem uma marcação que indique onde está localizada a DAM. Para isto, um dos radiologistas parceiros realizou a anotação da região de interesse que continha a DAM. Sendo assim, todas as imagens utilizadas neste trabalho foram manualmente anotadas por um radiologista com mais de 15 anos de experiência em mamografia. Na Figura 24, são apresentadas duas dessas imagens após a marcação.

Figura 24: Imagens clínicas após a marcação do radiologista.

Fonte: O autor, 2019.

4.2 Métodos Propostos

De acordo com os objetivos estabelecidos no Capítulo 1, os métodos propostos nesta tese de doutorado, para melhorar o desempenho dos sistemas de detecção automática da distorção arquitetural em imagens mamográficas 2D e 3D, serão apresentados nas

próximas seções. Para um melhor entendimento de onde cada abordagem se encaixa dentro do *pipeline* para a detecção de DAM a Figura 8 (pág. 37), apresentada anteriormente, será reapresentada com os devidos destaques.

4.2.1 Uso de Descritores de Micro-Padrões Local para a Discriminação de Regiões com Distorção Arquitetural

Devido às particularidades da DAM, inexistência de borda bem definida nem massa associada, existe uma grande dificuldade de selecioná-la dentre os mais diversos padrões de tecido sadio (SULEIMAN et al., 2016). Sendo assim, normalmente são empregados descritores de textura para fazer a descrição da áreas suspeitas de forma eficiente (ANAND; RATHANA, 2013; RANGAYYAN et al., 2013; KAMRA et al., 2016; MINA; ISA, 2015)

Como alguns trabalhos buscaram avaliar a aplicação dos descritores de Haralick, que se baseiam nas matrizes de coocorrência, o estudo (OLIVEIRA et al., 2016) (incluído na Seção 7.1) propôs outra abordagem: a utilização de descritores de micro-padrões local como o LBP e o LMP. Vale ressaltar que, este último nunca foi utilizado para esse fim. A nova abordagem, se encaixa na etapa referente a área sombreada em laranja da Figura 25.

Figura 25: Localização da proposta #1 dentro do *pipeline* para a detecção automática da DAM.

Fonte: O autor, 2019.

No momento do desenvolvimento deste trabalho, ainda não havia sido feito parceria com qualquer centro de imagem, sendo assim, as utilizadas foram da base DDSM. Um total de 200 imagens, todas com DAM, foram selecionadas. Em cada uma dessas imagens duas ROIs, de 128×180 pixels, foram recortadas. Uma ROI foi retirada da região na qual se encontrava a DAM a outra foi recortada de outra região qualquer da mama.

A avaliação do potencial do LMP para a descrição de regiões com DAM se deu com a comparação desde descritores populares na literatura, o LBP e os descritores de Haralick (Energia, Contraste, Correlação e Homogeneidade) tratados individualmente e em conjunto (chamado de Haralick-4) compondo um vetor de 4 números.

Após gerar cada vetor descritor, o processo de classificação se deu com a rede *perceptron*, com a qual várias topologias (de 1 a 30 neurônios) foram avaliadas. Cada topologia foi treinada 10 vezes de modo a evitar que a inicialização dos pesos aconteça em um ponto ruim do espaço de soluções. A escolha da melhor topologia utilizou como base a acurácia da classificação, no entanto, outras métricas foram calculadas para uma melhor avaliação.

Uma versão mais completa desse trabalho buscou avaliar o LMP não apenas na distinção entre DAM e tecido normal, mas também dentre outras anomalias (OLIVEIRA et al., 2017). Avaliando conjuntamente com os descritores comentados anteriormente, neste momento o objetivo era consolidar o LMP como um bom descritor mesmo para outras lesões. Mais detalhes podem ser encontrados no artigo em anexo, Seção 7.2.

4.2.2 Redução de Falsos-Positivos

Falsos-positivos representam um grande problema em todo *pipeline* para a detecção automática de anomalias em imagens médicas, não obstante essa situação se aplica àqueles que buscam detectar DAM (AYRES; RANGAYYAN, 2007; TANG et al., 2009; MORRA; DELSANTO; GRAZZINI, 2018).

Como foi comentado na Seção 3.4, o *pipeline* automático que tem sido adotado para a detecção de DAM apresenta uma grande limitação na etapa de redução de falsos positivos: o fato de se ter um parâmetro fixo para o *threshold* das regiões suspeitas. Supondo que este valor seja 50 isso implica que, se em uma imagem foram encontradas 100 regiões suspeitas, após serem ranqueadas com base na magnitude da região, apenas metade delas serão mantidas. Por sua vez, isso resulta em 50 regiões que deverão ser descartadas nas etapas seguintes do processo.

Figura 26: Localização da proposta #2 dentro do *pipeline* para a detecção automática da DAM.

Fonte: O autor, 2019.

O trabalho publicado na conferência *SPIE Medical Imaging 2018* (OLIVEIRA et al., 2018), anexado na Seção 7.3, teve por objetivo propor novos métodos para análise, e

descarte, de regiões suspeitas de serem DAM. Dentro do *pipeline* apresentado anteriormente, no Capítulo 3, a localização exata dessa proposta é destacada na Figura 26.

Essencialmente a análise de regiões suspeitas foi executada em outras imagens ao invés daquela que possui a magnitude da resposta do filtro de Gabor. Novamente os descritores de textura LBP e LMP foram utilizados juntamente com o CS-LBP, uma variação do LBP(HEIKKILÄ; PIETIKÄINEN; SCHMID, 2006). A aplicação de cada descritor em uma imagem irá gerar uma nova imagem (mapa de textura), o qual pode ser analisado. Sendo assim, além da proposta da análise de regiões suspeitas em outras imagens, outros 11 métodos, além do valor máximo, foram avaliados para fazer o ranqueamento.

Figura 27: Proposta para o descarte eficiente de falsos-positivos.

Fonte: Adaptado de (OLIVEIRA et al., 2018), 2019.

Todo o processo reproduzido na Figura 27, composto pela combinação de 5 tipos de imagens (intensidade, magnitude do filtro de Gabor, LMP, LBP e CS-LBP) com 11 métricas estatísticas, foi testado em 74 imagens de mamografia digital diagnosticadas com DAM. A aplicação de diferentes valores de *threshold* permitiu a análise da taxa de sucesso da abordagem aplicada.

4.2.3 Detecção de Distorção Arquitetural em Imagens de Tomossíntese Mamária

A detecção da DAM em imagens de tomossíntese digital da mama ainda é um problema que carece de soluções eficientes (SULEIMAN et al., 2016). O que se tem sugerido é a aplicação da mesma abordagem em cada fatia, o que está longe do ideal devido a quantidade enorme de regiões suspeitas e, consequentemente falsos-positivos (PALMA; BLOCH; MULLER, 2014; MORRA; COCCIA; CERQUITELLI, 2018; MORRA; DELSANTO; GRAZZINI, 2018).

Para se ter ideia, se for considerada uma mama de 5cm comprimida (o que resultaria em 50 imagens de fatias) e, uma média de 60 regiões suspeitas por imagem/fatia

após a aplicação do *pipeline* da Figura 8, a quantidade total de regiões suspeitas por mama/incidência seria de aproximadamente 3000. Regiões estas que deverão ser descritas e posteriormente classificadas.

Com o objetivo de propor uma solução para a detecção automática de DAM em imagens de DBT, o trabalho (OLIVEIRA et al., 2019) (reproduzido em sua totalidade na Seção 7.4) apresenta uma nova abordagem. Este trabalho, encaixa-se dentro da área sombreada da Figura 28, apesar do *pipeline* ali apresentado ser para mamografia digital. Utilizando um conjunto de 37 volumes de DBT, as etapas dessa proposta são resumidas na Figura 29.

Figura 28: Localização da proposta #3 dentro do *pipeline* para a detecção automática da DAM.

Fonte: O autor, 2019.

No processamento das fatias de DBT são empregados praticamente os mesmos procedimentos já comentados anteriormente no Capítulo 3, porém com algumas poucas variações. No entanto, a grande novidade está nos três últimos passos do fluxograma da Figura 29, que visam a maneira como cada região suspeita é tratada nas diferentes fatias.

Inicialmente, todas as fatias são processadas, como se fossem mamografias individuais. Com a obtenção das diferentes regiões suspeitas em cada fatia, essas regiões são conectadas às suas correspondentes nas fatias adjacentes. Como mostra a Figura 30, uma região suspeita na fatia #0, passando pela #8, é conectada com a sua respectiva nas N fatias seguintes. Essa conexão, gera uma trilha (em vermelho na Figura 30) que corresponde ao prolongamento da região suspeita nos diferentes níveis de profundidade da mama. A principal motivação por trás dessa abordagem é que, ao analisar um exame de DBT, o médico radiologista varre todas as fatias analisando as mudanças que ocorrem no tecido mamário.

A “conexão” de uma região suspeita nas N fatias adjacentes é feita com o uso da ferramenta TrackMate, um *plugin* do *software* Fiji/ImageJ (SCHINDELIN et al., 2012; TINEVEZ et al., 2017). O TrackMate foi inicialmente desenvolvido para o rastreamento de células em vídeos de microscopia. Desse modo, dado um conjunto de parâmetros

(descritos em (OLIVEIRA et al., 2019) e na Seção 7.4), ele é capaz de encontrar padrões comportamentais de um *frame* para outro do vídeo de modo a rastrear uma mesma célula que se deslocou. Esse tipo de abordagem, mesmo que a DBT não se trate de um vídeo, vai ao encontro do modo como o radiologista analisa o exame.

Figura 29: Proposta de um *pipeline* para a detecção de DAM em imagens de DBT.

Fonte: Adaptado de (OLIVEIRA et al., 2019), 2019.

Com a utilização do TrackMate foi possível realizar uma preliminar redução de falsos-positivos. Isso acontece porque existem regiões suspeitas que aparecem em uma única fatia, então assume-se que essa região é um FP, e portanto, trilhas com 1 ou dois pontos são descartadas juntamente com suas respectivas regiões suspeitas. Por outro lado, existem casos de uma mesma região suspeita se prolongar por 10 fatias, por exemplo, porém na fatia 5 não existe uma região que possa ser conectada, desse modo o TrackMate é capaz de realizar essa conexão pois fica claro que a região naquela fatia é um falso-negativo.

As trilhas remanescentes são utilizadas para a descrição de cada uma com base

nas suas ROIs. Para esse processo foi extraída uma ROI de $100 \times 100px$, da imagem de magnitudes do filtro de Gabor, tendo como centro as coordenadas do ponto da trilha. O uso de métricas estatísticas para a descrição dessas ROIs possibilita uma boa separação das trilhas que corresponde a DAM daquelas trilhas que são FPs, ou seja, correspondem aquelas regiões com tecido normal.

Figura 30: Exemplo de trilha (em vermelho) montada a partir da conexão de diferentes regiões suspeitas em cada fatia.

Fonte: O autor, 2019.

4.2.4 Uso de Arquitetura de Aprendizagem Profunda para o Reconhecimento de Áreas com DAM

Todos os descritores sejam eles de textura ou não, são desenvolvidos para um propósito específico. Eles são arquitetados de maneira que consigam descrever bem um determinado padrão, para isso são propostos todos os anos mais e mais descritores e variações dos já existentes. Diante desse cenário, fica a pergunta: e se fosse possível fazer um descritor específico para DAM?

Devido à variabilidade de uma DAM, seja no tamanho e/ou na intensidade com a qual a lesão aparece, fica difícil a arquitetura de um descritor específico para esse fim. Uma possível solução para esta situação vem de uma área relativamente nova, chamada *Deep Learning* ou Aprendizagem Profunda, que está ganhando espaço inclusive no ramo das imagens médicas (GREENSPAN; GINNEKEN; SUMMERS, 2016; LITJENS et al., 2017; MAYO; LEUNG, 2018; LAKHANI et al., 2018; AFSHAR et al., 2019).

Dentro da vasta gama de ferramentas que a *Deep Learning* proporciona, a classe de métodos chamados de *autoencoders* é uma delas (GUO et al., 2017; MAYO; LEUNG, 2018; AFSHAR et al., 2019). Neste, é possível treinar um modelo, capaz de aprender a descrever as características das imagens apresentadas a ele.

Baseando-se nos *autoencoders*, uma nova proposta apresentada em (OLIVEIRA et al., 2019) e anexado na Seção 7.5, busca apresentar esta técnica como uma tentativa de se obter um descritor que pode ser treinado para uma tarefa específica: descrever regiões com DAM. Dentro do processo clássico para detecção de DAM, apresentado novamente na Figura 31, esta nova proposta se localiza na área sombreada.

Figura 31: Localização da proposta #4 dentro do *pipeline* para a detecção automática da DAM.

Fonte: O autor, 2019.

O modelo de *autoencoder* proposto (Figura 32) é composto por duas partes, o *encoder* e o *decoder*, representado na figura pelos blocos em azul escuro e claro, respectivamente. Este tem como objetivo aprender a descrição das imagens de entrada, neste caso ROIs com DAM. Com uso de camadas convolucionais, blocos em azul escuro, a imagem é decomposta de maneira que a sua área sofra redução até atingir o centro do modelo, bloco em amarelo. Neste momento, o volume de $16 \times 16 \times 16$ passa por processos de-convolução de modo que a imagem de saída seja a mesma da entrada.

Por se tratar de uma arquitetura “profunda”, existe a necessidade de uma grande quantidade de imagens, demanda esta que dificulta o seu uso em imagens médicas. A solução é realizar o *data augmentation* (aumento dos dados), no qual as imagens sofrem transformações simples, como rotação e espelhamento, gerando assim novas amostras. Nesta proposta, a base de imagens é composta de 350 imagens de mamografia digital, 175/175 com/sem DAM. Para cada amostra (ROI), por meio de técnicas de *data augmentation*, foi possível gerar outras 400 amostras. Assim, a base de imagens final é composta de 140.000 amostras.

O processo de treinamento de um *autoencoder* busca minimizar o erro quadrático médio entre as imagens de entrada e as respectivas saídas. Assim, se garante que o código,

volume que se encontra no bloco amarelo, seja preciso o suficiente para que o processo inverso (decodificação) possa reconstruir a imagem original.

Figura 32: *Autoencoder* proposto para discriminação de regiões com DAM.

Fonte: Adaptado de (OLIVEIRA et al., 2019), 2019.

Todo esse procedimento comentado anteriormente serve para que o *encoder* aprenda a gerar um código preciso. Com o processo de treinamento do *autoencoder* finalizado, a parte referente ao *encoder* é removida e conectada a uma camada *flatten* (para vetorização do volume) e por fim a um classificador, neste caso uma rede neural *perceptron* multicamadas, como mostra a Figura 33.

Figura 33: Modelo *encoder*, retirado do *autoencoder* apresentado na Figura 32, conectado a uma rede neural.

Fonte: Adaptado de (OLIVEIRA et al., 2019), 2019.

Com o novo modelo (*encoder* e classificador) montado, é necessário um novo processo de treinamento, porém agora a otimização ocorrerá apenas nas camadas neurais, para que esta aprenda a classificar as amostras de acordo com a saída y esperada. O objetivo final desse novo modelo é que, ao ser apresentada uma nova imagem ela será inicialmente descrita pelo *encoder*, gerando um vetor de características e em seguida, classificada.

Resultados e Discussão

Nesta seção serão apresentados os resultados das 4 propostas comentadas anteriormente, no Capítulo 4. Vale ressaltar que essas propostas foram desenvolvidas em momentos diferentes e, portanto, as bases de imagens são diferentes para cada uma.

5.1 Uso de Descritores de Micro-Padrões Local para a Discriminação de Regiões com Distorção Arquitetural

Como foi reportado em (OLIVEIRA et al., 2016), com as imagens da base DDSM após a execução de 10 treinamentos de 30 diferentes topologias *perceptron* (de 1 a 30 neurônios), o treinamento que obteve a melhor acurácia para cada descritor foi recuperado. A Tabela 2, apresenta os principais resultados por meio das métricas acurácia, sensibilidade e AUC ROC. Para maiores detalhes recomenda-se ver a publicação original na Seção 7.1.

Tabela 2: Resultados das melhores topologias de rede neural para cada descritor utilizando a acurácia como critério de seleção.

Descritor	Acurácia	Sensibilidade	AUC
Contraste	90%	87%	0,94
Correlação	88%	86%	0,90
Energia	87%	91%	0,89
Homogeneidade	90%	94%	0,89
Haralick-4	93%	93%	0,96
LBP	85%	84%	0,87
LMP (0,32)	95%	97%	0,97

Fonte: Adaptado de (OLIVEIRA et al., 2016), 2019.

É possível observar que o LMP (com $\beta = 0,32$) obteve uma melhor acurácia, seguido pelo descritor Haralick-4 (a composição do contraste, correlação, energia e homogeneidade) com 93%. Apesar da pequena diferença, ao analisar a sensibilidade (taxa de detecção dos verdadeiros positivos), o LMP se saiu novamente superior aos demais. O LBP, que em vários trabalhos tem sido apontado como um dos descritores com um desempenho superior aos descritores de Haralick, não se saiu muito bem nesta avaliação, obtendo o pior desempenho dentre os descritores avaliados.

Figura 34: Curvas ROC calculadas utilizando os descritores LMP, LBP e Haralick-4.

Fonte: Adaptado de (OLIVEIRA et al., 2016), 2019.

Devido a Haralick-4 ter obtido um melhor resultado (AUC de 0,9576) dentre os descritores de Haralick, sua curva ROC foi colocada no gráfico da Figura 34, juntamente com aquelas dos descritores LMP e LBP. Em ordem decrescente de AUC, LMP é seguido por Haralick-4 e LBP. O LMP, apesar de ter apresentado a melhor AUC dentre os três (0,9689), este não se destaca pois Haralick-4 (0,9576) obteve um resultado muito aproximado.

A grande surpresa veio do desempenho baixo do LBP, visto que este tem se saído robusto em outras áreas. Por outro lado, ao se voltar para o problema que se deseja resolver: discriminar regiões com DAM dentre aquelas com tecido normal, é possível entender esse baixo desempenho. O LBP utiliza a função degrau como função de mapeamento, função esta que não é muito sensível para representar regiões onde ocorre uma pequena diferença

no contraste. Essas regiões com baixo contraste, são justamente aquelas onde se encontra a DAM devido ao seu alto grau de sutileza.

Como o LMP foi o descritor que obteve o melhor resultado, em um novo estudo (OLIVEIRA et al., 2017) este foi novamente avaliado sob as mesmas condições anteriores, no entanto, agora foram utilizadas imagens de mamografia digital.

A Tabela 3 apresenta os resultados das melhores topologias, de acordo com a acurácia. Para essa nova proposta o LMP (com $\beta = 0,2150$) obteve um melhor resultado em termos da AUC (0,9733), com acurácia de 90%. A energia obteve uma acurácia de 96% e AUC de 0,9464. Apesar do LMP gerar um vetor de 256 números e a energia ser apenas 1 valor, ambos descritores obtiveram os melhores resultados com uma rede *perceptron* de apenas 3 neurônios.

Tabela 3: Resultado da avaliação dos descritores de textura para discriminação de regiões com DAM em imagens de mamografia digital.

Descritor	Acurácia	AUC	Neurônios
LMP ($\beta = 0,2150$)	90%	0,9733	3
LBP	76%	0,7268	22
Haralick-4	86%	0,8201	10
Contraste	86%	0,8611	24
Correlação	90%	0,9156	2
Energia	96%	0,9464	3
Homogeneidade	86%	0,8611	8

Fonte: Adaptado de (OLIVEIRA et al., 2017), 2019.

As curvas ROC proveniente dos resultados com os descritores de Haralick, são apresentadas na Figura 35. Observa-se que, apesar do LMP possuir a maior AUC, a energia passa a ter um desempenho maior para uma mesma taxa de falsos-positivos próximo de 0. No entanto, a partir de uma taxa de FP de aproximadamente 0,15 o LMP obtém o melhor resultado dos três.

Com o objetivo de verificar o poder de discriminação dos descritores além da DAM, uma nova avaliação foi feita. Porém, nesta os descritores teriam que discriminar dentre regiões não apenas com DAM, mas também com nódulos, calcificações e tecido normal. Na Tabela 4 são apresentados os resultados dessa classificação, agora multi-classes, com a respectiva quantidade de neurônios necessária para o classificador.

É possível observar na Tabela 4 que, mesmo com outras áreas a serem discriminadas, o LMP ainda assim obteve o melhor resultado (acurácia média de 83%), seguido pelo

Figura 35: Curvas ROC resultante da discriminação de DAM e tecido normal com os descritores Energia, LMP e LBP, em imagens mamografia digital.

Fonte: (OLIVEIRA et al., 2017).

Haralick-4 e LBP. Dentre os três, LMP e Haralick-4 precisaram de uma rede com maior número de neurônios (30) para poder atingir o desempenho mencionado.

No gráfico da Figura 36 estão ilustrados os erros de classificação cometidos pelo classificador com base em cada um dos três descritores mais bem sucedidos. Neste gráfico, quanto mais próximo do centro melhor o resultado. Sendo assim, é notável que o LMP errou (acima de 0,6) bastante na distinção entre regiões de nódulos com calcificações. Além disso, houve uma taxa de 0,4 ao confundir regiões normais com calcificações. Vale ressaltar que o LMP obteve uma boa taxa de classificação nas situações que envolviam regiões com DAM, no gráfico observa-se que os menores picos são relacionados a essa anomalia.

O descritor Haralick-4, que obteve uma acurácia média de 82% (1% abaixo do LMP), é o que menos errou, pois seu único pico está no ponto de ter confundido regiões de DAM com regiões onde haviam nódulos. No entanto, outros picos menores podem ser visualizados dentro da área com o círculo cinza.

Novamente o LBP não se saiu bem, obtendo uma acurácia média de apenas 78%. No gráfico (Figura 36) observa-se vários picos representando as taxas de classificação erradas: nódulos com calcificações, nódulos com DAM, tecido normal com DAM, tecido

Tabela 4: Resultados em termos da acurácia média para a tarefa de distinção entre regiões com nódulos, calcificações, DAM e tecido normal.

Descritor	Acurácia média	Neurônios
LMP ($\beta = 0,5650$)	83%	30
LBP	78%	16
Haralick-4	82%	30
Contrast	76%	17
Correlation	75%	13
Energy	76%	26
Homogeneity	74%	11

Fonte: Adaptado de (OLIVEIRA et al., 2017), 2019.

normal com calcificações, dentre outros picos menores.

Figura 36: Representação das taxas de erros de classificação dos descritores LMP, LBP e Haralick-4.

Fonte: Adaptado de (OLIVEIRA et al., 2017), 2019.

Os resultados apresentados aqui, referente aos estudos (OLIVEIRA et al., 2016) e

(OLIVEIRA et al., 2017), mostraram que o LMP pode ser considerado um bom descritor para a discriminação de regiões com DAM. Em várias situações foi possível observar o seu desempenho superior se comparado com descritores clássicos como os de Haralick e o LBP. No entanto, apesar da sensibilidade em representar regiões de serem difíceis de serem mapeadas, o LMP apresenta um problema: o ajuste do parâmetro β . Como foi comentado anteriormente, na Seção 3.5.1.1, o β define a inclinação da curva *Sigmoid*, em outras palavras, esse parâmetro ajusta a sensibilidade do descritor. Achar o β ótimo para uma determinada tarefa, seja ela qual for, de forma automática ainda é um problema em aberto.

5.2 Redução de Falsos-Positivos

A eliminação de FPs é uma tarefa complicada e necessária para todo *pipeline* que visa a detecção de qualquer anomalia. Sendo assim, o estudo (OLIVEIRA et al., 2018) propôs várias alternativas para essa tarefa, visando o uso de um *threshold* que reduza a taxa de FPs sem que ocorra a eliminação da região com DAM.

Dado um conjunto de 74 imagens de mamografia digital com DAM, inicialmente foi verificada a quantidade média de FPs por imagem para que 100% das regiões com DAM fossem mantidas após o *threshold*, utilizando a abordagem usual. Essa análise, apresentada no gráfico da Figura 37, mostrou que 85 seria o *threshold* mínimo a ser utilizado e que no final da sua operação, haveria uma média de 44,4 regiões suspeitas por imagem.

Figura 37: Taxa de sucesso com base no processo usual, que utiliza o cálculo da média na imagem de magnitudes do filtro de Gabor, para análise e remoção de FPs.

Fonte: Adaptado de (OLIVEIRA et al., 2018), 2019.

No entanto, só o fato de mudar a métrica de ranqueamento, que é a média, para outras medidas estatísticas, já é possível obter resultados melhores como mostra o gráfico da Figura 38. Dentre as 11 métricas propostas, o *range* e a variância já conseguiram reduzir para 41,7 a média de regiões suspeitas por imagem com um *threshold* de 60.

Figura 38: Taxa de sucesso de outras 11 métricas estatísticas usando ainda a imagem de magnitudes de Gabor.

Fonte: (OLIVEIRA et al., 2018).

Na Tabela 5 está a relação das 5 imagens utilizadas nesta avaliação juntamente com a abordagem usual (sombreado em cinza) para comparação. Com o uso de outras imagens, ou mapas de textura, é possível obter resultados muito melhores do que aquele obtido ao utilizar a imagem resultante do filtro de Gabor. Nesta mesma imagem, como já foi comentado anteriormente, o cálculo da média já se mostra não ser uma boa métrica.

Dentre as imagens avaliadas, o mapa de textura do LMP produziu o melhor resultado possibilitando um *threshold* de 50 com uma taxa de 35,5 regiões suspeitas por imagem. No gráfico da Figura 39, observa-se o desempenho de todas as métricas aplicadas no mapa de textura gerado pelo LMP.

Como pôde ser observado, as métricas propostas anteriormente se saíram muito melhor do que a abordagem que é utilizada atualmente. Apesar de simples, no conjunto de 74 mamografias digitais avaliadas, houve uma grande redução na quantidade de regiões suspeitas por imagem, de 44,4 para 35,5. Esse tipo de seleção não supervisionada é de extrema importância para que não recaia sobre o descritor e, conseqüentemente o classificador, a carga de seleção das regiões de interesse, reduzindo assim o custo computacional.

Tabela 5: Imagens utilizadas com as respectivas métricas que obtiveram a menor quantidade de regiões suspeitas média. Em sombreado está a abordagem usualmente utilizada (RANGAYYAN et al., 2013; CASTI et al., 2016).

Imagem/Tex.	Métrica	Threshold	Regiões susp./Imagem
Gabor mag.	Média	85	44,4
Gabor mag.	Range/Variância	60	41,7
Intensidade	Contraste	55	36,8
LBP	<i>Skewness</i>	55	36,8
CS-LBP	Contraste	60	37,8
LMP ($\beta = 0,2$)	Correlação	50	35,5

Fonte: Adaptado de (OLIVEIRA et al., 2018).

Figura 39: Taxa de sucesso de outras 11 métricas estatísticas usando o mapa de Texturas gerado pelo LMP.

Fonte: Adaptado de (OLIVEIRA et al., 2018).

5.3 Detecção de Distorção Arquitetural em Imagens de Tomossíntese Mamária

Esta proposta teve por objetivo apresentar todo um *pipeline* para a detecção de DAM em imagens de DBT. Após a aplicação da sequência de passos apresentada na Seção 4.2.3, incluindo a classificação das trilhas remanescentes, foi possível montar a curva ROC apresentada na Figura 40. Nela observa-se uma baixa taxa de falsos-positivos

por volume mamário para cada ponto operacional de sensibilidade. Por exemplo, a uma sensibilidade de 0,9, houve uma média de 3,38 falsos positivos por volume.

Figura 40: Curva *Free ROC* referente ao processo de classificação das trilhas a nível de ROIs. Em destaque estão os valores de FP por volume de acordo com a sensibilidade.

Fonte: Adaptado de (OLIVEIRA et al., 2019), 2019.

Acredita-se que o bom desempenho do método proposto se deve principalmente ao modo como as trilhas são montadas, o que já possibilita a eliminação de vários FPs. Além disso, o processo de extração de características colaborativa, considerando as várias ROIs de uma mesma trilha, permite uma melhor descrição e discriminação daquelas ROIs, e consequentemente a trilha, corresponde a uma região com DAM ou não.

Na Figura 41 são apresentadas as trilhas, nas vistas lateral e superior, de um mesmo exame na incidência MLO. As trilhas em azuis são aquelas que foram corretamente classificadas como não DAM, a trilha vermelha representa um FP, enquanto que a verde o verdadeiro-positivo (VP). A região circulada em branco representa a marcação do radiologista onde se encontra a DAM.

Podem existir situações onde uma trilha está muito próxima da área marcada como DAM, como ilustra as imagens da Figura 42. A trilha em azul claro estava muito próxima daquela corretamente classificada como DAM e, parte dos seus pontos estavam contidos dentro da área que compreende a lesão. Para este caso, consideramos esta trilha como sendo um verdadeiro positivo já que é difícil mensurar a extensão de uma DAM. No entanto, ela foi classificada como não DAM, portanto essa trilha acontece de ser um FN.

As imagens da Figura 43 apresentam um exame na incidência CC com uma

Figura 41: Exame de DBT na incidência MLO, com as respectivas trilhas após a classificação sendo uma FP, na qual a DAM é representada por duas trilhas distintas.

Fonte: O autor, 2019.

Figura 42: Exame de DBT na incidência MLO, com uma trilha (em azul claro) muito próxima da DAM.

Fonte: O autor, 2019.

quantidade bastante reduzida de regiões suspeitas. Neste, a DAM é representada por um feixe de trilhas muito próximas, classificadas corretamente, no entanto outra trilha foi erroneamente definida como DAM (FP). Neste exame, o mais interessante é como a DAM

está representada no comportamento das trilhas envolvidas, o que é muito provável que em várias fatias mais de uma região suspeita culminou em representar a lesão.

Figura 43: Exame de DBT na incidência CC, com as respectivas trilhas após a classificação sendo uma FP.

Fonte: O autor, 2019.

A proposta apresentada aqui, tem obtido bons resultados para a discriminação de regiões com DAM. O processo de conexão das regiões suspeitas na direção Z do exame, em diferentes fatias, tem possibilitado uma maior acurácia do resultado. Isso mostra o quanto importante é a análise de uma região em fatias adjacentes, pois com isso, além de melhor caracterizar a DAM também ajuda a eliminar FPs. É necessário observar que foram testados diversos algoritmos que tentam descrever trilhas ao analisar a sua trajetória, como por exemplo o *Mean-squared Displacement*, dentre outros (MEIJERING; SMAL; DANUSER, 2006; GORELIK; GAUTREAU, 2014). Tais algoritmos não produziram um resultado satisfatório e portanto foram descartados.

5.4 Uso de Arquitetura de Aprendizagem Profunda para o Reconhecimento de Áreas com DAM

Como mencionado anteriormente (Seção 4.2.4), a proposta de uma arquitetura *autoencoder* buscou implementar um descritor que pudesse ser treinado para descrever exclusivamente áreas de DAM. O *dataset* apresentado (*Dataset #1*) foi arquitetado e “aumentado” para treinamento exclusivo do *autoencoder*. Após o seu treinamento e a junção do *encoder* com um classificador, este novo modelo foi novamente treinado com o mesmo *Dataset #1*.

Os resultados referentes a este treinamento são apresentados na primeira parte da Tabela 6. Para verificar a eficácia da proposta, descritores que anteriormente obtiveram bons resultados (como o LBP e a Energia) foram igualmente utilizados, junto com uma rede neural de mesma arquitetura, e treinados com o *Dataset #1*. As respectivas curvas ROC são apresentadas na Figura 44a.

Tabela 6: Resultados referente aos dois *datasets* avaliados. Os melhores resultados de cada coluna se encontram em negrito.

	<i>Dataset #1</i>		<i>Dataset #2</i>	
Descritor	Acurácia	AUC	Acurácia	AUC
Encoder	0,7052	0,7706	0,7535	0,8295
Energia	0,6908	0,7387	0,7473	0,8024
LBP	0,6471	0,7007	0,5000	0,5814

Fonte: (OLIVEIRA et al., 2019).

Uma segunda avaliação foi executada em um novo conjunto de imagens *Dataset #2*. Neste momento, o objetivo não era treinar qualquer modelo e, sim verificar o poder descritivo de cada um diante de um novo *dataset* que nenhum deles “viu”. Os resultados com este *dataset* são apresentados nas duas colunas à direita da Tabela 6 cujas curvas ROC estão na Figura 44b.

Figura 44: Curvas ROC referente aos *datasets* utilizados para treinamento e predição.

Fonte: Adaptador de (OLIVEIRA et al., 2019), 2019.

Os resultados com o *Dataset #2* talvez sejam os mais relevantes, pois mostram o quão poderoso é o conjunto (descriptor + classificador). Dentre os três descritores o *encoder* obteve o melhor resultado, em termos de acurácia, com 0,7535, seguido pela energia com 0,7473 e o LBP com o pior resultado 0,5000. O cálculo das áreas sob a curva (AUC) ROC seguem o mesmo padrão da acurácia.

Com os valores aproximados do *encoder* e a energia, não é possível dizer com propriedade qual é o melhor dos dois. Até onde se conhece, o grande problema de um modelo de aprendizagem profunda está na sua necessidade de um grande conjunto de dados para o treinamento. Isso nos leva a crer que, o *autoencoder* tem grande potencial para resultados melhores, porém com uma maior quantidade de imagens.

Vale ressaltar que, apesar do LMP ter apresentado um resultado superior nas propostas anteriores, ele não foi utilizado nessa avaliação devido a dificuldade de treinamento por conta do parâmetro β já discutido anteriormente. Como o treinamento feito com o *Dataset #1* levaria muito tempo para ser executado pois acarretaria a execução de N treinamentos, sendo N a quantidade de valores β a serem avaliados, isso inviabilizou o teste no momento em que o trabalho estava sendo desenvolvido.

Conclusões e Trabalhos Futuros

A Distorção Arquitetural Mamária (DAM) é uma anomalia que por muitas vezes não é detectada devido a sua sutileza. Apesar da dificuldade em encontrá-la durante a análise de uma imagem mamográfica, essa anomalia deve ser diagnosticada se ali estiver presente. Sua detecção seguida do tratamento apropriado, poderá contribuir para a efetividade do tratamento aumentando as chances de que seja barrado o desenvolvimento de um possível câncer.

Como o rastreamento da doença é feito normalmente com a radiografia das mamas, a mamografia digital é o exame mais utilizado no mundo. No entanto, um exame mais poderoso vem sendo utilizado em grandes centros desde 2011: a Tomossíntese Digital Mamária (DBT, sigla em inglês). Espera-se que em um futuro próximo a DBT venha a substituir totalmente a mamografia digital, isso devido às inúmeras vantagens que tal exame traz.

Devido a importância de detectar a DAM é de extrema necessidade o desenvolvimento de métodos que possam detectá-la para auxiliar os médicos radiologistas nessa tarefa. Mesmo com a utilização de exames como DBT, a detecção de DAM ainda é insatisfatória (SULEIMAN et al., 2016). Os pontos reforçados anteriormente justificam o desenvolvimento de estudos como este, principalmente quando se trata de exames de DBT, onde existe uma carência de métodos para a detecção dessa anomalia.

Com base nos trabalhos propostos até o presente momento é possível notar que, apesar das dificuldades acerca da detecção da DAM, aos poucos o objetivo vem sendo atingido. No entanto muito ainda precisa ser feito, seja para sua detecção em mamografia digital ou em imagens de DBT.

Este trabalho de doutorado teve por objetivo propor quatro abordagens para a detecção de DAM sendo que, três delas foram validadas na mamografia digital e uma, desenvolvida especialmente para DBT. A seguir, essas propostas são brevemente comentadas além de sugestões de continuação do trabalho em cada uma dessas frentes:

Proposta 1: como a caracterização de regiões com DAM é uma tarefa complexa, o uso de descritores de textura é recorrente. Esta primeira proposta teve por objetivo apresentar um novo descritor para esse fim, o LMP. Os testes reportados nos trabalhos (OLIVEIRA et al., 2016; OLIVEIRA et al., 2017) (seções 7.1 e 7.2), mostraram que o LMP é muito mais sensível do que os descritores mais populares da literatura, os descritores de Haralick e o LBP. Foi possível observar que o LMP supera esses descritores com sucesso na tarefa de discriminar áreas com DAM daquelas que correspondem a tecido normal. Durante o desenvolvimento dessa proposta foi possível comprovar a dificuldade de se utilizar o LMP, pois esse possui um parâmetro (β) que é fundamental para a sua eficiência. Ainda não existe método determinístico para o cálculo desse parâmetro, portanto Algoritmo Genético, *Particle Swarm Optimization*, e até mesmo força-bruta tem sido empregada para a sua determinação de acordo com a tarefa. Outros descritores vêm sendo propostos a cada dia, bem como métodos para a efetiva “engenharia das características” como métodos de redução de dimensionalidade e fusão de descritores, dentre outros. Talvez, já exista algum método capaz de descrever regiões com sucesso, porém este ainda não foi avaliado para detectar a DAM;

Proposta 2: em um *pipeline* automático para a detecção da DAM em mamografia digital existe o momento onde, após a análise do tecido mamário, é necessário eliminar o máximo de regiões suspeitas, sem que haja a perda daquela que corresponde a DAM. Esta segunda proposta (OLIVEIRA et al., 2018) (Seção 7.3) teve por objetivo apresentar e avaliar novos meios para a eliminação desses falsos-positivos (FPs). Como essa tarefa é algo não supervisionado, pois naquele momento não se sabe o que é tecido sadio ou lesionado, foi proposto o uso de diversas outras métricas estatísticas para a ordenação e corte da lista de regiões suspeitas. Além das medidas estatísticas, a utilização de outras imagens diferente da abordagem usual que utiliza a imagem de magnitudes do filtro de Gabor, foi apresentada. Mapas de textura produzidos por descritores como o LMP, LBP e CS-LBP, mostraram possuir propriedades interessantes para o cálculo das métricas. Assim, foi comprovado que é possível reduzir com eficácia a quantidade de FPs sem que a região com DAM seja perdida no processo. A maior limitação desse trabalho está no fato de que ainda assim é necessária a determinação de um valor de *threshold* para o descarte de regiões suspeitas. Do ponto em que se encontra esta frente é necessária uma análise inicial do conjunto de imagens a ser processado para que seja determinado um *threshold* ótimo. Nesse sentido, uma nova proposta que busque a automatização desse valor, definindo um *threshold* ótimo por imagem, seria de grande importância;

Proposta 3: a detecção da DAM em imagens de DBT, assim como na mamografia digital, é um problema em aberto. O grande “problema” do desenvolvimento de métodos que utilizem DBT está no fato de que este exame gera um volume de dados muito maior

do que o seu antecessor. Sendo assim, a terceira proposta (OLIVEIRA et al., 2019) (Seção 7.4) tem por objetivo a apresentação de um *pipeline* totalmente automático para a detecção de DAM em imagens de DBT. Nos passos iniciais do processo são utilizadas as mesmas ferramentas empregadas em mamografia digital, nesse caso em cada fatia de tomossíntese é aplicado o mesmo conjunto de procedimentos. A maior novidade dessa proposta está no modo como os FPs são eliminados, já que em cada fatia existe uma grande quantidade deles. Com o uso de uma ferramenta para o rastreamento de células em vídeos de microscopia, foi possível “conectar” as regiões suspeitas em uma fatia com as mesmas regiões nas fatias adjacentes, gerando assim trilhas que conectam a mesma região em diferentes imagens. Com isso, foi possível eliminar uma grande quantidade de regiões suspeitas com base nas trilhas formadas como por exemplo, aquelas que possuíam apenas um ou dois pontos ou aquelas que atravessavam a mama em toda a sua longitude. As trilhas remanescentes tiveram suas regiões descritas de modo que pudessem ser eliminadas após a classificação. A descrição das trilhas com base nas regiões que as compõem foi feita de maneira que houvesse a contribuição de cada região no cálculo de métricas que pudesse discriminar aquela determinada trilha como sendo normal ou DAM. Longe de ser um método soberano para a detecção de DAM em DBT, esta proposta apresenta como um dos pontos fracos a quantidade de processamento necessário nas etapas iniciais do *pipeline*. Como foi utilizado como base o mesmo procedimento para o uso em imagens de mamografia, várias tarefas são executadas várias vezes, a mesma quantidade de fatias produzidas no exame. Desse modo, a ampliação dessa proposta com o uso de métodos que atuem em volume de dados tridimensionais, tem muito a contribuir. Apesar de ainda ser um tópico emergente, o uso de descritores 3D, bem como a convolução tridimensional, esta é uma área que necessita de muito estudo e desenvolvimento já que a cada mês surgem computadores cada vez mais potentes;

Proposta 4: arquiteturas profundas de aprendizagem (*Deep Learning* em inglês) é um tópico que vem ganhando visibilidade nos últimos anos, alavancado principalmente pelo uso das unidades gráficas de processamento, as GPUs. Dentre os mais diversos modelos que a *deep learning* apresenta, estão aqueles chamados de *autoencoders*. Os *autoencoders* são arquiteturas que podem aprender como descrever os dados de entrada. A grande vantagem desse tipo de modelo é a possibilidade de, após treinado, utilizá-lo como um descritor especializado para descrever o tipo de imagem no qual foi treinado. Nesse sentido, essa última proposta (OLIVEIRA et al., 2019) (Seção 7.5) busca apresentar um modelo de *autoencoder* capaz de descrever regiões com DAM. Após o processo de treinamento com 140 mil amostras (geradas por meio da execução do *data augmentation*) o *encoder* (a primeira parte de um *autoencoder*) foi conectado a um classificador, de modo que este possa discriminar entre regiões com DAM e normais com base no vetor descritor gerado pelo *encoder*. Uma nova etapa de

treinamento deve ser feita, neste momento apenas para o ajuste do classificador. Nessa proposta o uso do *autoencoder* foi comparado com descritores de textura apresentados anteriormente, comprovando assim o grande potencial dessa abordagem para a detecção da DAM. Quando se usa uma arquitetura profunda essencialmente duas coisas são necessárias, um grande volume de dados e poder computacional. O primeiro requerimento passa a ser um problema quando o objetivo é a detecção da DAM, visto que mesmo os grandes centros possuem poucos, talvez não mais que mil casos de DAM reportados por ano. Assim, o treinamento de uma rede que necessite de milhões de imagem para que atinja uma taxa de convergência razoável parece impossível. O uso de técnicas de *data augmentation* e *transfer learning* podem minimizar esse problema, porém ainda assim não se resolve por completo. O desenvolvimento de novas técnicas de *data augmentation* podem vir a contribuir significativamente nesse quesito. No entanto, a aplicação de outras técnicas de *deep learning* com novas arquiteturas parecem sem mais promissoras, mas mesmo assim, ainda se recai no limite dos dados.

Todas as propostas aqui apresentadas foram testadas em imagens clínicas de pacientes reais para uma avaliação realista do que se propõe. A detecção da DAM é uma tarefa bastante desafiadora e necessita de mais atenção da comunidade científica, vista a grande vantagem de se ter o diagnóstico precoce de um possível câncer de mama. Longe de ser a palavra final quando o assunto é detecção de DAM, este trabalho de doutorado buscou, acima de tudo, contribuir para que esta importante lesão possa ser diagnosticada com eficiência e precisão.

Publicações Oriundas da Tese

Neste capítulo são apresentados, em ordem cronológica de publicação, os artigos oriundos desta tese de doutorado.

7.1 Artigo 1: XII Workshop de Visão Computacional

Trabalho publicado e apresentado no congresso nacional
XII Workshop de Visão Computacional (WVC 2016).

Detection of Architectural Distortion in Digitized Screen-film Mammograms Using Texture Descriptors

Helder C. R. de Oliveira, Diego R. Moraes, Adilson Gonzaga, Marcelo A. C. Vieira

Escola de Engenharia de São Carlos (EESC)

Departamento de Engenharia Elétrica e de Computação

Universidade de São Paulo (USP)

{heldercro, diego.moraes}@usp.br, {agonzaga, mvieira}@sc.usp.br

Abstract—This paper proposes a computer-based system for automatic detection of architectural distortion of the breast (ADB) in digitized screen-film mammograms. ADB is a subtle contraction of the breast tissue that may represent the earliest sign of breast cancer. The proposed system has essentially two steps: feature extraction and classification. For the first step we evaluated six different texture descriptors: Local Mapped Pattern (LMP), Local Binary Pattern (LBP) and four Haralick's descriptors. The second step was performed with a multilayer perceptron (MLP). The best topology was chosen based on the classification accuracy by varying the amount of neurons in the hidden layer from 1 to 30. We evaluate the performance of the proposed detection scheme using a dataset of 400 regions of interest (ROIs) extracted from the mammograms. Results in terms of the receiver operating characteristics curve (ROC) have shown that the LMP descriptor performed better than the similar descriptors.

Index Terms—architectural distortion; mammography; local mapped pattern; local binary pattern; haralick; texture descriptor.

I. INTRODUÇÃO

O câncer de mama é a principal doença causadora de morte entre as mulheres no Brasil e em todo o mundo [1]. Em 2012, ela foi o responsável pelo óbito de aproximadamente 522.000 mulheres. Estima-se que, a cada ano, há o surgimento de 1,67 milhão de novos casos. No Brasil, dados do INCA (Instituto Nacional de Câncer) indicam que este é o segundo tipo de câncer de maior incidência. Para o ano de 2016, estima-se o aparecimento de aproximadamente 57 mil novos casos no Brasil, dos quais aproximadamente 24% resultarão em morte [2].

Quando detectado em seu estágio inicial, a probabilidade de cura do câncer de mama aumenta em aproximadamente 30%, já que o tratamento da doença em estágio avançado é muito mais difícil [3], [4]. Para a descoberta prematura dessa doença a mamografia é o exame mais utilizado em todo o mundo [5], [6]. Durante a análise de uma imagem mamográfica, o radiologista busca por sinais que podem indicar um câncer em formação como: nódulos; microcalcificações e; distorção arquitetural.

A Distorção Arquitetural Mamária (DAM) é uma contração sutil do tecido mamário que pode representar o sinal mais

precoce de um câncer de mama em formação, surgindo antes mesmo da formação de nódulos e microcalcificações [7], [8]. Estudos recentes mostraram que se a DAM for detectada precocemente, o prognóstico no tratamento do câncer é muito superior do que no caso da detecção precoce de microcalcificações e nódulos [7]. No entanto, a DAM é de difícil detecção nos exames de mamografia convencional, pois não possui borda bem definida e não ocorre a formação de massa visível. Dessa forma, mesmo o radiologista mais experiente pode não ser capaz de identificá-la [9].

Quando detectada, a DAM normalmente é classificada na categoria BI-RADS 4C [10], pois se trata de um achado suspeito que pode ser o início de um câncer maligno. Nesse caso, apenas a biópsia pode descartar a hipótese de malignidade [10]. Além disso, a DAM é o achado mais comum em exames retrospectivos de mamografia quando ocorrem casos de falso-negativo, ou seja, quando o médico não detectou a lesão que já estava presente [7].

Com o intuito de auxiliar os radiologistas no diagnóstico precoce do câncer de mama, sistemas computacionais de auxílio ao diagnóstico (*Computer Aided Detection* - CAD) são normalmente utilizados na prática clínica. Esses softwares têm por objetivo indicar, de modo automático, regiões suspeitas na mamografia, fornecendo uma “segunda opinião” ao médico radiologista [11], [12]. Diversos estudos têm mostrado que os radiologistas, quando assistidos pelo CAD, são capazes de detectar na mamografia aproximadamente 90% das microcalcificações e nódulos. No entanto, para o caso de detecção da DAM, o desempenho dos sistemas CAD ainda é insatisfatório, com taxa de detecção abaixo dos 50% [13]–[15].

Assim, o objetivo deste trabalho é propor um sistema automatizado para a detecção de distorção arquitetural mamária. O sistema consiste basicamente de duas etapas: extração de características e classificação. Para a etapa de extração de características, foram avaliados 6 descritores de textura: o *Local Binary Pattern* (LBP) [16], o *Local Mapped Pattern* (LMP) [17], [18] e quatro descritores de Haralick (Contraste, Correlação, Energia e Homogeneidade) [19]. Para a etapa de classificação, foi avaliado o uso de uma rede neural do tipo Perceptron Multi Camadas (PMC), contendo apenas uma

camada escondida. Várias topologias de rede neural foram avaliadas onde a diferença entre elas é a quantidade de neurônios na camada escondida que, nesse caso, variou-se de 1 a 30 neurônios. A escolha da melhor topologia para cada descritor foi feita com base no cálculo da acurácia.

II. MATERIAIS E MÉTODOS

A. Detecção de Distorção Arquitetural Mamária

Com a premissa de que exista um classificador treinado, de maneira geral os métodos desenvolvidos para a detecção de distorção arquitetural mamária em imagens de mamografia seguem a sequência ilustrada na Figura 1.

Figura 1. Sequência de etapas comumente abordadas em um detector de distorção arquitetural.

A sequência de passos apresentada anteriormente, na Figura 1, é brevemente comentada a seguir:

- 1) O processo inicia-se com uma imagem adquirida por meio de exame mamográfico. A detecção das áreas suspeitas (*Region of Interest* - ROI) tem por objetivo analisar a arquitetura da mama em cada ponto da imagem, buscando regiões onde aparentemente houve uma possível distorção de tal estrutura;
- 2) Todas as ROIs são então submetidas a aplicação dos extratores de características (detectores de textura) selecionados para essa abordagem;
- 3) Os descritores produzidos pelos métodos de extração de características são utilizados como entrada em um classificador previamente treinado. Com base na resposta do classificador, almeja-se a eliminação de todos os ROIs que são Falsos Positivos (*False Positive* - FP), ou seja, regiões que foram selecionadas como suspeitas mas que não possuem distorção arquitetural;
- 4) Por fim, espera-se que restem apenas as regiões que são Verdadeiros Positivos (*True Positive* - TP), regiões que efetivamente possuem a anomalia a ser detectada.

Para efeito de simplicidade, foi assumido que as etapas iniciais que visam o levantamento das áreas suspeitas foram realizadas com sucesso. Para o presente trabalho foi investigado um conjunto de 400 ROIs, ou seja, pequenas imagens que representam as regiões consideradas suspeitas.

B. Imagens de Mamografia Digitalizadas

Uma vez que o objetivo deste trabalho é propor um conjunto de métodos para a detecção de distorção arquitetural mamária,

se faz então necessária a utilização de imagens clínicas de pacientes que possuam laudo médico para a validação. Foram utilizadas imagens do banco de imagens *Digital Database for Screening Mammography* [20] (DDSM), consagrado na literatura e disponibilizado publicamente na internet.

O DDSM [20] é um banco de imagens fruto do esforço conjunto das seguintes instituições americanas: *University of South Florida*, *Sandia National Laboratories* e *Washington University School of Medicine*. Apesar de as imagens serem de exames de mamografia analógica (tela-filme) digitalizada, onde ocorre uma significativa perda de qualidade, tais dados ainda são muito utilizados em diversos trabalhos propostos na literatura atual [14], [21]–[26]. Na Figura 2, são apresentadas duas imagens do DDSM utilizadas neste trabalho.

Figura 2. Imagens do banco de imagens DDSM [20] nas views (a) Média Lateral Oblíqua (MLO) e (b) Crânio-caudal (CC).

Todas as imagens do DDSM que possuem alguma anomalia são acompanhadas do laudo médico que, dentre outras informações, indica o tipo de anomalia e as coordenadas em que ela se encontra. Em posse dessas informações, foram selecionadas 200 imagens cujos laudos comprovavam a existência da distorção arquitetural mamária.

Cada imagem foi recortada manualmente em duas ROIs, uma contendo a distorção arquitetural e outra contendo tecido normal, sadio. As regiões selecionadas passaram por uma etapa de pré-processamento, onde foi efetuado o realce do contraste e a sub-amostragem deixando as ROIs com dimensão 128×128 pixels. O primeiro processamento tem por objetivo tornar mais visível a distorção arquitetural bem como a estrutura mamária, enquanto que o segundo visa a atenuação de qualquer ruído presente e manter uma uniformidade nos tamanhos das ROIs.

Tais pré-processamentos são perfeitamente válidos de serem feitos pois, durante a análise da imagem para o diagnóstico, o software utilizado pelo médico permite o acesso a esse ferramental. Na Figura 3, é apresentado parte do conjunto de

ROIs utilizadas sendo, (a-c) com distorção arquitetural e (b-d) considerados normais.

Figura 3. Amostras após o pré-processamento, (a-c) com distorção arquitetural e (b-d) normais.

C. Descritores de Textura

Após a seleção e pré-processamento das ROIs, foram aplicados para cada uma das amostras os descritores de características. Neste trabalho, foram utilizados quatro extratores de características propostos por Haralick [19], que se baseiam no cálculo das matrizes de co-ocorrência, ao verificar o pixel à direita com distância 1:

$$\text{Contraste} = \sum_i^k \sum_j^k (i - j)^2 p_{ij} \quad (1)$$

$$\text{Correlação} = \sum_i^k \sum_j^k \frac{(i - m_r)(j - m_c)p_{ij}}{\sigma_r \sigma_c} \quad (2)$$

$$\text{Energia} = \sum_i^k \sum_j^k p_{ij}^2 \quad (3)$$

$$\text{Homogeneidade} = \sum_i^k \sum_j^k \frac{p_{ij}}{1 + |i - j|} \quad (4)$$

nas quais, i e j representam coordenadas da matriz de co-ocorrência de dimensão k , p_{ij} é o ij -ésimo elemento da matriz e $\sigma_r \neq 0$ e $\sigma_c \neq 0$ representam os desvios-padrão nas linhas e colunas, respectivamente.

Outro descritor utilizado foi o *Local Binary Pattern* (LBP), em sua formulação clássica proposto por Ojala, Pietikainen e Maenpää no ano de 1996 [16]. O LBP busca caracterizar a textura local de uma imagem utilizando para isso um operador,

apresentado na Equação 5, que irá analisar os tons de cinza da vizinhança P do pixel central g_0 :

$$LBP = \sum_{i=1}^P S(g_0, g_i) 2^{i-1} \quad (5)$$

na qual $S(\cdot)$ é a função de *Heavyside*, definida por:

$$S(g_0, g_i) = f_g = \begin{cases} 1, & \text{se } g_i \geq g_0 \\ 0, & \text{caso contrário.} \end{cases} \quad (6)$$

Após o cálculo do LBP em cada pixel da imagem, tem-se então um mapa de LBPs cujos valores são utilizados para a composição do histograma de 256 valores, definindo assim o vetor de características. Vale ressaltar que o operador *LBP* é invariante com relação à escala de tons de cinza e à rotação.

Uma proposta de expandir o LBP para vizinhança genérica de $W \times W$ pixels e mapear o resultado para um histograma de B bins foi apresentada em [17]. A nova formulação, chamada de *Local Mapped Pattern* (LMP), é apresentada na Equação 7, na qual f_g é a função de *Heavyside* mostrada na Equação 6, B o número de bins do histograma e $P(i)$ a matriz de pesos binários dada pela Equação 8.

$$LBP = \text{round} \left(\frac{\sum_{i=1}^W f_g \cdot P(i)}{\sum_{i=1}^W P(i)} (B - 1) \right) \quad (7)$$

$$P(i) = \begin{bmatrix} 1 & 2 & 4 \\ 128 & 0 & 8 \\ 64 & 32 & 16 \end{bmatrix} \quad (8)$$

Além das mudanças apresentadas anteriormente, o LMP apresentou a possibilidade de alteração da função de mapeamento de *Heavyside* para *Sigmoidal*, que é uma suavização da função degrau apresentada na Equação 6:

$$f_g = \frac{1}{1 + e^{-\frac{(g_i - g_0)}{\beta}}} \quad (9)$$

Com a função *Sigmoidal* é possível sintonizar o parâmetro β , que define a inclinação da curva, criando vários descritores. Tal recurso tornar possível a criação de um descritor otimizado para a aplicação desejada. Levando em conta as considerações apresentadas anteriormente, a formulação completa do LMP é apresentada na Equação 10.

$$LMP = \text{round} \left(\frac{\sum_{i=1}^W \left(\frac{1}{1 + e^{-\frac{(g_i - g_0)}{\beta}}} \cdot P(i) \right)}{\sum_{i=1}^W P(i)} (B - 1) \right) \quad (10)$$

D. Classificador

Após a extração de características das ROIs, os descritores produzidos foram utilizados como entrada para uma rede neural artificial com arquitetura do tipo Perceptron Multi Camadas (PMC), contendo apenas uma camada escondida.

A escolha do PMC se deve ao fato de ser o classificador mais utilizado para o reconhecimento de padrões [27]–[29]. Como saída, a rede neural deverá retornar os valores: 1 se houver distorção arquitetural e 0 caso contrário.

A implementação do PMC utilizada é a disponível no MATLAB R2013b, devido a sua facilidade de prototipagem e realização de simulações. A rede foi treinada com o algoritmo *backpropagation* com gradiente escalado. O máximo de épocas foi fixado em 500 e com um erro máximo de 10^{-100} .

As 400 amostras foram divididas de forma aleatória, metade de uma classe e metade de outra: sendo 70% utilizadas para treinamento, 15% para validação e; 15% para teste. O conjunto para treinamento foi utilizado para o ajuste dos pesos da rede neural. No entanto, com o objetivo de evitar o *overfitting* (situação onde a rede “memoriza” os dados de treinamento e erra quando um novo conjunto de dados é apresentado) o cálculo do erro é baseado em um novo conjunto de dados, por isso se faz necessário o conjunto de validação [27], [28]. É necessário comentar que os resultados reportados neste artigo foram baseados apenas no grupo de teste.

Uma rede PMC pode ser montada com diferentes topologias, com uma ou mais camadas ocultas e quantos neurônios forem necessários em cada camada. No presente trabalho, a rede utilizada é graficamente representada na Figura 4.

Figura 4. Exemplo da rede neural utilizada, na qual: x_M representa a quantidade de valores do vetor de características; N é a quantidade de neurônios na camada escondida e; y a saída da rede.

Como pôde ser observado na Figura 4, a rede PMC com apenas uma camada oculta tem os seguintes parâmetros: x_M entradas, de acordo com o tamanho do vetor de características; N neurônios na camada oculta, pois foram avaliadas as topologias de 1 a 30 neurônios e; a saída binária y , que indicará 1 se a imagem possui distorção arquitetural ou 0 se não houver.

No processo de treinamento da rede neural, os pesos sinápticos foram inicializados com valores aleatórios entre 0 e 1. Para evitar viés nos resultados, foram realizados 10 treinamentos para cada uma das 30 topologias candidatas (de 1 a 30 neurônios), ou seja, foram executados 300 treinamentos para assim identificar a melhor topologia com melhor treinamento.

III. RESULTADOS E DISCUSSÃO

A melhor topologia selecionada foi aquela na qual determinado treinamento obteve a melhor acurácia. Foram avaliados os extratores de características: LBP, LMP e Haralick. Para o caso do método proposto por Haralick, que se baseia no cálculo da matriz de co-ocorrência, foram avaliadas as seguintes características juntas e individualmente: Contraste, Correlação, Energia e Homogeneidade.

Os resultados da melhor topologia para cada descritor são apresentados na Tabela I, na qual, além da acurácia são apresentados os cálculos da Sensibilidade, Especificidade, Precisão, F1 Score e a Área Sob a Curva (*Area Under Curve* - AUC) *Receiver Operating Characteristic* (ROC) [30].

Tabela I
RESULTADOS DAS MELHORES TOPOLOGIAS DE REDE NEURAL PARA CADA DESCRIPTOR UTILIZANDO A ACURÁCIA COMO CRITÉRIO DE SELEÇÃO.

Descritor	Acurácia	Sensibilidade	Especificidade	Precisão	F1 Score	AUC
Contraste	90%	87%	93%	93%	90%	0.94
Correlação	88%	86%	92%	94%	89%	0.90
Energia	87%	91%	82%	85%	88%	0.89
Homogeneidade	90%	94%	84%	88%	91%	0.89
Haralick-4	93%	93%	94%	93%	93%	0.96
LBP	85%	84%	86%	91%	88%	0.87
LMP (0,32)	95%	97%	93%	93%	95%	0.97

Como é possível observar na tabela anterior, em negrito estão destacados o melhor valor de cada métrica. Os resultados mostram que o LMP (com parâmetro $\beta = 0,32$) possui a melhor taxa de detecção para as quatro métricas: Acurácia, Sensibilidade, F1 Score e AUC. Além disso, observa-se que o Contraste, Energia, Homogeneidade e o LBP, não obtiveram o melhor resultado em nenhuma das métricas avaliadas.

A curva ROC apresenta a relação entre a taxa de verdadeiros positivos (sensitividade) e falsos positivos (especificidade). As curvas ROC das características relacionadas na Tabela I, são apresentadas no conjunto de imagens da Figura 5.

As curvas ROC da Figura 5a apresentam resultados obtidos para os descritores de Haralick analisados separadamente: Contraste, Correlação, Energia e a Homogeneidade; e também foi analisado o desempenho quando todas essas características foram consideradas em conjunto, sendo que esse agrupamento foi chamado de “Haralick-4”. Sendo assim, na Figura 5a, pode ser observado que Haralick-4 possui um melhor desempenho, $AUC = 0,9576$, se for comparado com cada uma das características separadas.

Como Haralick-4 se mostrou o melhor dentre todas as avaliações com os descritores separados, este foi comparado com os descritores LMP e LBP, e o resultado é apresentado na Figura 5b. Pode-se observar que o LMP apresenta resultados superiores aos outros descritores, com, $AUC = 0,9689$. Além disso, se tomarmos 10% de falsos positivos (valor 0,1 no eixo X da curva ROC apresentada na Figura 5b) é possível

Figura 5. Curvas ROC calculadas utilizando (a) descritores de Haralick e (b) LMP, LBP e Haralick-4).

visualizar que o LBP possui uma taxa de Verdadeiros Positivos (VP) um pouco maior que 80%, seguido por Haralick-4 com aproximadamente 90%, enquanto que o LMP apresenta uma taxa de 94%.

IV. CONCLUSÕES

Este estudo apresentou uma proposta de sistema automático para a detecção de distorção arquitetural mamária. Como essa tarefa engloba diversas etapas, neste trabalho foram abordadas apenas aquelas relacionadas à extração de características e classificação. Para a extração de características, foram avaliados seis métodos de descrição de textura, sendo dois deles (LMP e LBP) baseados em micro-padrões locais enquanto que os outros quatro são baseados nas matrizes de co-ocorrência propostas por Haralick. Para a classificação, foi utilizada uma rede PMC com uma única camada escondida, na qual foram avaliadas várias topologias com neurônios variando de 1 a 30.

Os resultados têm mostrado que o LMP se mostrou superior aos demais na maioria das métricas avaliadas, possuindo uma acurácia de 95% e $AUC = 0,9689$. O LMP apresentou inclusive um poder de discriminação superior em relação aos

outros métodos que são largamente utilizados na literatura com a finalidade de detectar distorção arquitetural ou para descrever texturas em imagens naturais.

Os resultados apresentados foram colhidos utilizando imagens de mamografia tela-filme digitalizadas que foram manualmente recortadas e pré-processadas. Novas etapas deste trabalho visam a utilização de imagens de mamografia digital (*Full Field Digital Mammography* - FFDm), bem como a implementação de todo o *pipeline* para a detecção de distorção arquitetural mamária apresentado na Figura 1.

AGRADECIMENTOS

Os autores gostariam de agradecer à FAPESP (Proc. 2015/20812-5) pelo apoio financeiro.

REFERÊNCIAS

- [1] OMS, "Organização mundial da saúde," 2016. [Online]. Available: <http://www.who.int/>
- [2] INCA, "Instituto nacional de câncer josé alencar gomes da silva," 2016. [Online]. Available: <http://www.inca.gov.br/>
- [3] J. G. Elmore, C. Y. Nakano, T. D. Koepsell, L. M. Desnick, C. J. D'Orsi, and D. F. Ransohoff, "International variation in screening mammography interpretations in community-based programs," *J. Natl. Cancer Inst.*, vol. 95, no. 18, pp. 1384–1393, 2003.
- [4] U. Veronesi, P. Boyle, A. Goldhirsch, R. Orecchia, and G. Viale, "Breast Cancer," *The Lancet*, vol. 365, pp. 1727–1741, 2005.
- [5] A. D. A. Maidment, "Digital mammography," *Seminars in Roentgenology*, vol. 38, no. 3, pp. 216–230, 2003.
- [6] C. G. Glynn, D. M. Farria, B. S. Monsees, J. T. Salzman, K. N. Wiele, and C. F. Hildebolt, "Effect of transition to digital mammography on clinical outcomes," *Radiology*, vol. 260, no. 3, pp. 664–670, 2011.
- [7] S. Gaur, V. Dialani, P. J. Slanetz, and R. L. Eisenberg, "Architectural Distortion of the Breast," *American Journal of Roentgenology*, vol. 201, pp. 662–670, 2013.
- [8] M. Bahl, J. A. Baker, E. N. Kinsey, and S. V. Ghate, "Architectural Distortion on Mammography: Correlation With Pathologic Outcomes and Predictors of Malignancy," *American Journal of Roentgenology*, vol. 205, no. 6, pp. 1339–1345, nov 2015.
- [9] K. M. Ray, E. Turner, E. A. Sickles, and B. N. Joe, "Suspicious Findings at Digital Breast Tomosynthesis Occult to Conventional Digital Mammography: Imaging Features and Pathology Findings," *The Breast Journal*, pp. 1–5, 2015.
- [10] C. J. D'Orsi, Ed., *ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System*. American College of Radiology, 2013.
- [11] J. C. Dean and C. C. Ilvento, "Improved cancer detection using computer-aided detection with diagnostic and screening mammography: Prospective study of 104 cancers," *American Journal of Roentgenology*, vol. 187, no. 1, pp. 20–28, 2006.
- [12] S. K. Yang, W. K. Moon, N. Cho, J. S. Park, J. H. Cha, S. M. Kim, S. J. Kim, and J.-G. Im, "Screening mammography-detected cancers: Sensitivity of a computer-aided detection system applied to full-field digital mammograms," *Radiology*, vol. 244, no. 1, pp. 104–111, 2007.
- [13] J. A. Baker, E. L. Rosen, J. Y. Lo, E. I. Gimenez, R. Walsh, and M. S. Soo, "Computer-aided detection (CAD) in screening mammography: Sensitivity of commercial CAD systems for detecting architectural distortion," *American Journal of Roentgenology*, vol. 181, no. 4, pp. 1083–1088, 2003.
- [14] R. M. Rangayyan, S. Banik, and J. E. L. Desautels, "Computer-Aided Detection of Architectural Distortion in Prior Mammograms of Interval Cancer," *Journal of Digital Imaging*, vol. 23, no. 5, pp. 611–631, 2010.
- [15] W. I. Suleiman, M. F. McEntee, S. J. Lewis, M. A. Rawashdeh, D. Georgian-Smith, R. Heard, K. Tapia, and P. C. Brennan, "In the digital era, architectural distortion remains a challenging radiological task," *Clinical Radiology*, vol. 71, no. 1, pp. e35–e40, 2016.
- [16] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, 1996.
- [17] R. T. Vieira, C. E. de Oliveira Chierici, C. T. Ferraz, and A. Gonzaga, *Local Fuzzy Pattern: A New Way for Micro-pattern Analysis*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 602–611.

- [18] C. T. Ferraz, O. Pereira, M. V. Rosa, and A. Gonzaga, "Object Recognition Based on Bag of Features and a New Local Pattern Descriptor," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 28, no. 08, p. 1455010, 2014.
- [19] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.
- [20] M. Heath, K. Bowyer, D. Kopans, R. Moore, and W. P. Kegelmeyer, "The digital database for screening mammography," in *Proceedings of the 5th international workshop on digital mammography*, 2000, pp. 212–218.
- [21] M. Jasionowska and A. Przelaskowski, "A two-step method for detection of architectural distortions in mammograms," *Information Technology in Biomedicine*, vol. 69, pp. 73–84, 2010.
- [22] S. Banik, R. M. Rangayyan, and J. E. L. Desautels, "Detection of Architectural Distortion in Prior Mammograms," vol. 30, no. 2, pp. 279–294, 2011.
- [23] T. Handa, X. Zhang, N. Homma, T. Ishibashi, Y. Kawasumi, M. Abe, N. Sugita, and M. Yoshizawa, "DoG-Based Detection of Architectural Distortion in Mammographic Images for Computer-Aided Detection," pp. 762–767, 2012.
- [24] O. Netprasat, S. Auephanwiriyakul, and N. Theera-Umpon, "Architectural Distortion Detection from Mammograms Using Support Vector Machine," in *International Joint Conference on Neural Networks*, 2014, pp. 3258–3264.
- [25] A. Kamra, V. K. Jain, S. Singh, and S. Mittal, "Characterization of Architectural Distortion in Mammograms Based on Texture Analysis Using Support Vector Machine Classifier with Clinical Evaluation," *Journal of Digital Imaging*, 2015.
- [26] I. Zyout and R. Togneri, "Empirical Mode Decomposition of Digital Mammograms for the Statistical based Characterization of Architectural Distortion," pp. 109–112, 2015.
- [27] B. D. Ripley, *Pattern Recognition and Neural Networks*. Cambridge University Press, 2007.
- [28] C. M. Bishop, *Neural Networks for Pattern Recognition*. Clarendon Press, 1995.
- [29] I. N. da Silva, D. H. Spatti, and R. A. Flauzino, *Redes Neurais Artificiais para Engenharia e Ciências Aplicadas*. Artliber, 2010.
- [30] D. M. W. Powers, "Evaluation: From precision, recall and f-measure to roc., informedness, markedness & correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.

7.2 Artigo 2: *SPIE Medical Imaging 2017*

Trabalho publicado e apresentado no congresso internacional
SPIE Medical Imaging (SPIE 2017),
conferência *Computer-Aided Diagnosis*:
<http://doi.org/10.1117/12.2255516>

A New Texture Descriptor Based on Local Micro-Pattern for Detection of Architectural Distortion in Mammographic Images

Helder C. R. de Oliveira^a, Diego R. Moraes^a, Gustavo A. Reche^a, Lucas R. Borges^a,
Juliana H. Catani^b, Nestor de Barros^b, Carlos F. E. Melo^c,
Adilson Gonzaga^a, and Marcelo A. C. Vieira^a

^aDepartment of Electrical Engineering, University of São Paulo, São Carlos, Brazil.

^bDepartment of Radiology, University of São Paulo, São Paulo, Brazil.

^cClínica Eco & Mama Diagnóstico Digital, São Carlos, Brazil.

{heldercro, diego.moraes, gustavo.reche, lucas.rodrigues.borges}@usp.br,
{juliana.catani, nestor.barros}@hc.fm.usp.br,
clinicaecoemama@hotmail.com, {agonzaga, mvieira}@sc.usp.br

ABSTRACT

This paper presents a new local micro-pattern texture descriptor for the detection of Architectural Distortion (AD) in digital mammography images. AD is a subtle contraction of breast parenchyma that may represent an early sign of breast cancer. Due to its subtlety and variability, AD is more difficult to detect compared to microcalcifications and masses, and is commonly found in retrospective evaluations of false-negative mammograms. Several computer-based systems have been proposed for automatic detection of AD, but their performance are still unsatisfactory. The proposed descriptor, Local Mapped Pattern (LMP), is a generalization of the Local Binary Pattern (LBP), which is considered one of the most powerful feature descriptor for texture classification in digital images. Compared to LBP, the LMP descriptor captures more effectively the minor differences between the local image pixels. Moreover, LMP is a parametric model which can be optimized for the desired application. In our work, the LMP performance was compared to the LBP and four Haralick's texture descriptors for the classification of 400 regions of interest (ROIs) extracted from clinical mammograms. ROIs were selected and divided into four classes: AD, normal tissue, microcalcifications and masses. Feature vectors were used as input to a multilayer perceptron neural network, with a single hidden layer. Results showed that LMP is a good descriptor to distinguish AD from other anomalies in digital mammography. LMP performance was slightly better than the LBP and comparable to Haralick's descriptors (mean classification accuracy = 83%).

Keywords: architectural distortion, mammography, local mapped pattern, local binary pattern, haralick, texture descriptor.

1. INTRODUCTION

Breast cancer is the most common cancer in women worldwide. It was responsible for 522,000 deaths in 2002,¹ and the estimate for the next years is around 1.67 million cases per year. To reduce the mortality rate, governments of several countries have adopted screening programs to detect breast cancer in its earliest stage. When detected in initial stage, the probability of cure of breast cancer increases in approximately 30%, since the treatment of this disease in advanced stage is more difficult.^{2,3}

Mammography is the most used exam for breast cancer screening in the world.^{4,5} While reading a mammogram, radiologists look for signs that could indicate the development of cancer. The most common signs are: Masses, Microcalcifications, Bilateral Asymmetry and Architectural Distortion (AD). Among them, AD is more difficult to detect because it is a subtle contraction of breast tissue.⁶

AD may be the earliest sign of breast cancer because it manifests before nodules and microcalcifications.^{7,8} Several studies have shown that if AD is early detected, the prognosis of the cancer treatment is superior if compared to cases where masses or microcalcifications were detected.^{7,9}

Medical Imaging 2017: Computer-Aided Diagnosis, edited by Samuel G. Armato III, Nicholas A. Petrick,
Proc. of SPIE Vol. 10134, 101342U · © 2017 SPIE · CCC code: 1605-7422/17/\$18 · doi: 10.1117/12.2255516

When detected, the AD is usually classified as BI-RADS 4C,¹⁰ because it is a suspicious finding that may represent a manifestation of malignant breast cancer. In this case, biopsy is often necessary to exclude the suspicious of malignancy.¹⁰ Moreover, AD is the most common finding in retrospective assessments of false-negative mammography.⁷

To assist radiologists in the early detection of breast cancer, Computer Aided Detection (CAD) systems are commonly used in clinical practice. These softwares are intended to automatically indicate suspicious regions in the mammogram, providing a “second opinion” to the radiologist.^{11,12} Several studies have shown that radiologists, when assisted by CAD, are capable of detecting approximately 90% of microcalcifications and masses on mammography. However, for the cases with architectural distortion, the performance of CAD systems is still unsatisfactory, with a detection rate around 50%.^{13–15}

Thus, the objective of this work is to present a new texture descriptor for the detection of architectural distortion in Full-Field Digital Mammography (FFDM) images. The proposed descriptor, the Local Mapped Pattern (LMP),^{16,17} is a generalization of Local Binary Pattern (LBP),¹⁸ which is widely used as descriptor for texture classification in digital images. LMP is a parametric model which can be optimized for the desired application. In this work, the LMP performance was compared to the LBP and four Haralick’s descriptors (Contrast, Correlation, Energy and Homogeneity)^{19,20} for classification of regions of interest extracted from clinical FFDM images. The classification stage was performed by using a Multilayer Perceptron (MLP) neural network with one hidden layer. Several neural network topologies were evaluated, changing the amount of neurons in the hidden layer from 1 to 30. The choice of the best topology for each descriptor was made based on the classification accuracy.

2. MATERIALS AND METHODS

2.1 Clinical Image Dataset

Figure 1. Example of images used on this work: (a) Digital mammogram with architectural distortion marked by a radiologist; (b) ROI with AD, extracted from the marked region; (c) ROI with normal tissue extracted from another region in the same image.

The dataset used for the evaluation of the proposed descriptor is composed of 300 clinical FFDM images. The images were obtained from the Institute of Radiology of São Paulo (*Instituto de Radiologia de São Paulo - InRAD*) at the University of São Paulo Medical School (*Faculdade de Medicina da Universidade de São Paulo*).

- FMUSP). We have got the institutional review board approval to use these images in this research (number #1.581.220 - *Certificado de Apresentação para Apreciação Ética* - CAAE #56699016.7.0000.0065).

The dataset was divided into three sets of 100 FFDM images, based on the most important finding described in the medical report. A radiologist marked the lesion in each mammogram and a Region of Interest (ROI) was manually extracted. To get a sample of healthy tissue, another ROI was extracted from each image in a region without any lesion (normal tissue). The Figure 1a shows a digital mammogram with AD marked by a radiologist and the ROI extracted from this region (Figure 1b). In addition, an ROI with normal tissue extracted from the same mammogram is shown in Figure 1c.

2.2 Feature Extractors

2.2.1 Local Texture Descriptors

Local descriptors are based on image features extracted from small regions of an image, using histograms to generate a feature vector. The Local Binary Pattern (LBP) is a widely used and well-reputed texture descriptor based on local micro-patterns for texture classification in digital images.^{17,18}

In its original formulation, the LBP considers the gray level of P pixels in the neighborhood of a central pixel g_0 , according to Equation 1, where $S(\cdot)$ is the HeavySide mapping function presented at Equation 2.

$$LBP = \sum_{i=1}^P S(g_0, g_i) 2^{i-1}, \quad (1)$$

where,

$$S(g_0, g_i) = f_g = \begin{cases} 1, & \text{if } g_i \geq g_0 \\ 0, & \text{otherwise.} \end{cases} \quad (2)$$

A proposal of expanding the LBP for a generic neighborhood of $W \times W$ pixels and matching the result for a histogram of B bins was presented in Refs. 16 and 17. This new formulation, called Local Mapped Pattern (LMP), is presented in Equation 3:

$$LMP = \text{round} \left(\frac{\sum_{i=1}^W f_g \cdot P(i)}{\sum_{i=1}^W P(i)} (B - 1) \right), \quad (3)$$

where f_g is the same mapping function of Equation 2, B is the number of bins of the histogram and, $P(i)$ is the binary weight matrix given by Equation 4:

$$P(i) = \begin{bmatrix} 1 & 2 & 4 \\ 128 & 0 & 8 \\ 64 & 32 & 16 \end{bmatrix} \quad (4)$$

Moreover, the LMP formulation presents a possibility of changing the mapping function from HeavySide to Sigmoidal, as shown in Equation 5, which is a smoothing version of the step function:

$$f_g = \frac{1}{1 + e^{\frac{-(g_i - g_0)}{\beta}}} \quad (5)$$

Using this formulation, it is possible to tune the β parameter, which defines the slope of the sigmoid curve, creating several different descriptors which can optimize the LMP for a specific application. Taking into account all considerations presented previously, the new formulation of LMP is shown in Equation 6:

$$LMP = \text{round} \left(\frac{\sum_{i=1}^W \left(\frac{1}{1 + e^{\frac{-(g_i - g_0)}{\beta}}} \cdot P(i) \right)}{\sum_{i=1}^W P(i)} (B - 1) \right) \quad (6)$$

It is worth to mention that if the parameter β of the mapping function is equal to 0, the LMP converges to the LBP descriptor.

2.2.2 Haralick's Co-occurrence Matrices

In addition to the LBP and LMP, we also evaluated Haralick's texture descriptors for the classification of the selected ROIs.^{19,20} The descriptors proposed by Haralick are based on the calculation of co-occurrence matrices. In this work were considered four Haralick's features: contrast, energy, correlation and homogeneity; defined by the following equations:

$$\text{Contrast} = \sum_i^k \sum_j^k (i - j)^2 p_{ij} \quad (7)$$

$$\text{Correlation} = \sum_i^k \sum_j^k \frac{(i - m_r)(j - m_c)p_{ij}}{\sigma_r \sigma_c} \quad (8)$$

$$\text{Energy} = \sum_i^k \sum_j^k p_{ij}^2 \quad (9)$$

$$\text{Homogeneity} = \sum_i^k \sum_j^k \frac{p_{ij}}{1 + |i - j|} \quad (10)$$

where, i and j are spatial coordinates of the co-occurrence matrix of order k , p_{ij} is the ij -element of the matrix and $\sigma_r \neq 0$ and $\sigma_c \neq 0$ are the standard deviations computed through its rows and columns, respectively.

2.3 Classification

A Multilayer Perceptron Neural Network (MLP-NN) with a single hidden layer was used as the classifier. The feature vectors used as input were generated considering all descriptors mentioned before: LMP, LBP and the four Haralick's descriptors. The best β parameter for the LMP was empirically defined based on the classification accuracy. Regarding the Haralick's features, we evaluated the performance of each feature separately and also all four descriptors combined.

The samples used for the MLP-NN were separated randomly in 70% for training, 15% for validation and 15% for test. All results reported here were based on test samples. To avoid classification bias, we evaluated 30 different MLP-NN topologies varying the number of neurons in the hidden layer from 1 to 30. Moreover, we performed 10 training for each topology and selected the best one for each feature descriptor, based on the measurement of classification accuracy.

As illustrated in the flowchart at Figure 2, two different experiments were conducted in order to assess the descriptors on the detection of several mammography findings. The first experiment evaluated the classification of 200 ROIs into two classes: AD and normal tissue. The second experiment was conducted to evaluate a multi-class classification of 400 ROIs into four different classes, according to the anomaly detected: AD, mass, microcalcification and normal tissue. The dataset was composed of 100 samples of each class.

For the binary classification performed in the first experiment, the best topology was chosen based on classification accuracy. Further analysis regarding the classifier performance was done based on a Receiver Operating

Characteristic (ROC) curve. For the multi-class classification performed in the second experiment, the best topology was chosen based on the mean accuracy, which takes into account the mean accuracy among all four classes.²¹ Furthermore, ROC analysis cannot be performed when the number of classes is greater than two. In this case, we used a *Cobweb* representation²² to visualize the multi-class classification performance. The *Cobweb* representation provides a quick way to visualize classifier performance. The general idea is to present a graphic representation of the confusion matrix generated by the classifier. Each column in the class confusion matrix is normalized by the total value of this class. The misclassification values are chosen from the confusion ratio matrix to form a *Cobweb* representation. Each corner of the polygon represents the normalized misclassification value which comes from the off-diagonal of the confusion ratio matrix.²²

Figure 2. Sequence of steps describing the present study.

3. RESULTS AND DISCUSSION

3.1 Binary Classification

Figure 3. ROC curves for AD *vs.* normal tissue classification using (a) four Haralick's features separately and combined; (b) using the LMP, LBP and the Energy feature.

Figure 3a shows the ROC curves calculated for classification of AD *vs.* normal tissue using each one of the

four Haralick's descriptors separately and combined (Haralick-4). Among all Haralick's descriptors, the Energy feature (continuous line) provided the best result in terms of Area Under the Curve (AUC). Moreover, Energy descriptor alone overcome the performance of all four Haralick's descriptors combined (dashed line).

Since Energy descriptor presented the best classification performance in terms of AUC, we compared this descriptor to the local texture descriptors LBP and LMP. The β parameter for LMP descriptor was tuned using a mixed-integer genetic algorithm,²³ in order to provide the best performance in terms of classification accuracy. Figure 3b shows the ROC curves of the Haralick's Energy feature, LBP and LMP descriptors. It is possible to verify that the performance of the LMP descriptor is higher than the LBP and comparable to the Haralick's features (AUC = 0.9733).

Table 1 shows the results of the classification accuracy, AUC and number of neurons in the hidden layer for the best MLP-NN topology selected for each descriptor. In terms of accuracy, Energy feature provided the best result: 96% using 8 neurons in the hidden layer followed by the Correlation feature: accuracy = 90% with two neurons. The LMP ($\beta = 0.2150$) provided 90% of accuracy using only 3 neurons in the hidden layer.

Table 1. Results of accuracy, AUC and the number of neurons used for AD detection using the best topology of a MLP neural network.

<i>Descriptor</i>	<i>Accuracy</i>	<i>AUC</i>	<i>Neurons</i>
LMP ($\beta = 0.2150$)	90%	0.9733	3
LBP	76%	0.7268	22
Haralick-4	86%	0.8201	10
Contrast	86%	0.8611	24
Correlation	90%	0.9156	2
Energy	96%	0.9464	3
Homogeneity	86%	0.8611	8

3.2 Multi-class Classification

The second experiment involved a multi-class classification problem. In this scenario we investigated the LMP ability, compared to others descriptors, to classify four different lesions in digital mammograms (AD, mass, microcalcification and normal tissue). Figures 4 and 5 show the *Cobweb* representation of our 4-class classifier performance. In a 4-class classifier, a 4×4 class confusion matrix was used to analyze errors. The values chosen from a 4-class classification confusion matrix form a twelve-dimensional misclassification point. Each edge of the radar plot shows a situation when a sample of "Class A" (True State) was classified as "Class B" (Decision State). For a chance classification, the misclassification value of the confusion ratio matrix is 0.25 for each class. It is represented by a gray dashed line in the graphs of figures 4 and 5. A point within the chance performance polygon shows a better performance than chance classifier.

In Figure 4, only the misclassifications based on Haralick's descriptors are presented. Correlation, Homogeneity and Energy features provided the worst classification performance, since they have the highest peaks in the graph. Particularly, the Correlation generated one unique significant peak which represents the situation where AD was erroneously classified as Mass. Energy feature performed the biggest misclassification in detecting Calcification as AD, where it missed all cases. Based on the *Cobweb* analysis, the best classification performance was reached when all four Haralick's descriptors were combined (Haralick-4). However, it provided a high misclassification rate in classifying AD as Mass. Because Haralick-4 provided the best results, it was selected to be compared to the local texture descriptors evaluated on this work (LBP and LMP).

Figure 5 shows the *Cobweb* representation for the classifier when using Haralick-4, LBP and LMP descriptors. The LBP descriptor generated several peaks of wrong classifications. The LMP descriptor generated two peaks of wrong classifications (Mass $\rightarrow$ Calc and Normal $\rightarrow$ Calc). However, all misclassifications regarding architectural distortion stayed within the "classification by chance" polygon.

Cobweb representation of misclassification (True state -> Decision state)

Figure 4. *Cobweb* representation for MLP neural network classifier using four Haralick's features separately and combined. The gray dashed line represents the "classification by chance" polygon.

Cobweb representation of misclassification (True state -> Decision state)

Figure 5. *Cobweb* representation for MLP neural network classifier using LMP, LBP and four Haralick's features combined. The gray dashed line represents the "classification by chance" polygon.

Table 2 shows the mean accuracy obtained for the 4-class classifier, as well as the amount of neurons in the hidden layer necessary to reach the best result in each case. LMP performed slightly better than the LBP

descriptor and comparable to all four Haralick's descriptor combined.

Table 2. Results of mean accuracy²¹ and the amount of neurons used for the multi-class classification experiment.

<i>Descriptor</i>	<i>Mean accuracy</i>	<i>Neurons</i>
LMP ($\beta = 0.5650$)	83%	30
LBP	78%	16
Haralick-4	82%	30
Contrast	76%	17
Correlation	75%	13
Energy	76%	26
Homogeneity	74%	11

4. CONCLUSIONS

In this work, we presented a new texture descriptor, Local Mapped Pattern (LMP), to improve the detection of architectural distortion in digital mammography. To make a fair comparison, the LMP was compared to the most used feature descriptors in the literature for texture discrimination: LBP and four Haralick's descriptors. For all tests, clinical images of full-field digital mammography were used.

Results showed that LMP is a good descriptor to distinguish AD from other anomalies in digital mammography. LMP performance was slightly better than the LBP and comparable to Haralick's descriptors. Future works are necessary to improve the LMP performance.

ACKNOWLEDGMENTS

This project is supported by São Paulo Research Foundation (FAPESP) grant #2015/20812–5 and the National Council for Scientific and Technological Development (CNPq).

REFERENCES

- [1] WHO, "World health organization - breast cancer." <http://www.who.int/>. (Accessed: 15 January 2017).
- [2] Elmore, J. G., Nakano, C. Y., Koepsell, T. D., Desnick, L. M., D'Orsi, C. J., and Ransohoff, D. F., "International variation in screening mammography interpretations in community-based programs," *J. Natl. Cancer Inst.* **95**(18), 1384–1393 (2003).
- [3] Veronesi, U., Boyle, P., Goldhirsch, A., Orecchia, R., and Viale, G., "Breast Cancer," *The Lancet* **365**, 1727–1741 (2005).
- [4] Maidment, A. D. A., "Digital mammography," *Seminars in Roentgenology* **38**(3), 216–230 (2003).
- [5] Glynn, C. G., Farria, D. M., Monsees, B. S., Salcman, J. T., Wiele, K. N., and Hildebolt, C. F., "Effect of transition to digital mammography on clinical outcomes," *Radiology* **260**(3), 664–670 (2011).
- [6] Banik, S., Rangayyan, R. M., and Desautels, J. E. L., "Detection of Architectural Distortion in Prior Mammograms," *IEEE Transactions on Medical Imaging* **30**(2), 279–294 (2011).
- [7] Gaur, S., Dialani, V., Slanetz, P. J., and Eisenberg, R. L., "Architectural Distortion of the Breast," *American Journal of Roentgenology* **201**, 662–670 (2013).
- [8] Bahl, M., Baker, J. A., Kinsey, E. N., and Ghatge, S. V., "Architectural Distortion on Mammography: Correlation With Pathologic Outcomes and Predictors of Malignancy," *American Journal of Roentgenology* **205**(6), 1339–1345 (2015).
- [9] Ray, K. M., Turner, E., Sickles, E. A., and Joe, B. N., "Suspicious Findings at Digital Breast Tomosynthesis Occult to Conventional Digital Mammography: Imaging Features and Pathology Findings," *The Breast Journal*, 1–5 (2015).

- [10] D'Orsi, C. J., ed., [*ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System*], American College of Radiology (2013).
- [11] Dean, J. C. and Ilvento, C. C., "Improved Cancer Detection Using Computer-Aided Detection with Diagnostic and Screening Mammography - prospective study of 104 cancers," *American Journal of Roentgenology* **187**, 20–28 (2006).
- [12] Yang, S. K., Moon, W. K., Cho, N., Park, J. S., Cha, J. H., Kim, S. M., Kim, S. J., and Im, J.-G., "Screening mammography-detected cancers: Sensitivity of a computer-aided detection system applied to full-field digital mammograms," *Radiology* **244**(1), 104–111 (2007).
- [13] Baker, J. A., Rosen, E. L., Lo, J. Y., Gimenez, E. I., Walsh, R., and Soo, M. S., "Computer-aided detection (CAD) in screening mammography: Sensitivity of commercial CAD systems for detecting architectural distortion," *American Journal of Roentgenology* **181**(4), 1083–1088 (2003).
- [14] Rangayyan, R. M., Banik, S., and Desautels, J. E. L., "Computer-Aided Detection of Architectural Distortion in Prior Mammograms of Interval Cancer," *Journal of Digital Imaging* **23**(5), 611–631 (2010).
- [15] Suleiman, W. I., McEntee, M. F., Lewis, S. J., Rawashdeh, M. A., Georgian-smith, D., Heard, R., Tapia, K., and Brennan, P. C., "In the digital era, architectural distortion remains a challenging radiological task," *Clinical Radiology* **71**(1), e35–e40 (2016).
- [16] Vieira, R. T., de Oliveira Chierici, C. E., Ferraz, C. T., and Gonzaga, A., [*Local Fuzzy Pattern: A New Way for Micro-pattern Analysis*], 602–611, Springer Berlin Heidelberg, Berlin, Heidelberg (2012).
- [17] Ferraz, C. T., Pereira, O., Rosa, M. V., and Gonzaga, A., "Object Recognition Based on Bag of Features and a New Local Pattern Descriptor," *International Journal of Pattern Recognition and Artificial Intelligence* **28**(08), 1455010 (2014).
- [18] Ojala, T., Pietikäinen, M., and Harwood, D., "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognition* **29**(1), 51–59 (1996).
- [19] Haralick, R. M., Shanmugam, K., and Dinstein, I., "Textural Features for Image Classification," *IEEE Transactions on Systems, Man, and Cybernetics* **SMC-3**(6), 610–621 (1973).
- [20] Haralick, R. M., "Statistical and structural approaches to texture," in [*Proceedings of the IEEE*], **67**(5), 786–804 (1979).
- [21] Sokolova, M. and Lapalme, G., "A systematic analysis of performance measures for classification tasks," *Information Processing and Management* **45**(4), 427–437 (2009).
- [22] Diri, B. and Albayrak, S., "Visualization and analysis of classifiers performance in multi-class medical data," *Expert Systems with Applications* **34**(1), 628–634 (2008).
- [23] Mitchell, M., [*An Introduction to Genetic Algorithms*], MIT Press, Cambridge, MA, USA (1996).

7.3 Artigo 3: *SPIE Medical Imaging 2018*

Trabalho apresentado e publicado no congresso internacional
SPIE Medical Imaging (SPIE 2018),
conferência *Computer-Aided Diagnosis*:
<http://doi.org/10.1117/12.2293388>

Reduction of false-positives in a CAD scheme for automated detection of architectural distortion in digital mammography

Helder C. R. de Oliveira^a, Arianna Mencattini^b, Paola Casti^b, Eugenio Martinelli^b,
Corrado di Natale^b, Juliana H. Catani^c, Nestor de Barros^c, Carlos F. E. Melo^d,
Adilson Gonzaga^a, and Marcelo A. C. Vieira^{a,*}

^aDept. of Electrical and Computing Engineering, University of São Paulo, São Carlos, Brazil

^bDept. of Electronic Engineering, University of Rome - Tor Vergata, Rome, Italy

^cDept. of Radiology, University of São Paulo, São Paulo, Brazil

^dClínica Eco & Mama Diagnóstico Digital, São Carlos, Brazil

heldercro@usp.br, {agonzaga, mvieira}@sc.usp.br

{mencattini, casti, martinelli, dinatale}@ing.uniroma2.it

{juliana.catani, nestor.barros}@hc.fm.usp.br, gerencia@ecoemama.com.br

ABSTRACT

This paper proposes a method to reduce the number of false-positives (FP) in a computer-aided detection (CAD) scheme for automated detection of architectural distortion (AD) in digital mammography. AD is a subtle contraction of breast parenchyma that may represent an early sign of breast cancer. Due to its subtlety and variability, AD is more difficult to detect compared to microcalcifications and masses, and is commonly found in retrospective evaluations of false-negative mammograms. Several computer-based systems have been proposed for automated detection of AD in breast images. The usual approach is automatically detect possible sites of AD in a mammographic image (segmentation step) and then use a classifier to eliminate the false-positives and identify the suspicious regions (classification step). This paper focus on the optimization of the segmentation step to reduce the number of FPs that is used as input to the classifier. The proposal is to use statistical measurements to score the segmented regions and then apply a threshold to select a small quantity of regions that should be submitted to the classification step, improving the detection performance of a CAD scheme. We evaluated 12 image features to score and select suspicious regions of 74 clinical Full-Field Digital Mammography (FFDM). All images in this dataset contained at least one region with AD previously marked by an expert radiologist. The results showed that the proposed method can reduce the false positives of the segmentation step of the CAD scheme from 43.4 false positives (FP) per image to 34.5 FP per image, without increasing the number of false negatives.

Keywords: architectural distortion, digital mammography, Gabor filter, local binary pattern, local mapped pattern, texture descriptor.

1. INTRODUCTION

Architectural distortion (AD) is a subtle contraction of breast tissue and can represent the earliest sign of breast cancer in its developing stage.^{1,2} Some studies have shown the importance of early detection of AD in breast images because the prognosis of the cancer treatment is superior if compared to cases where masses or microcalcifications were detected first.^{2,3} Moreover, AD is the most common finding in retrospective evaluation of false-negative mammograms.⁴

Computer-Aided Detection (CAD) systems are commonly used in clinical practice to assist radiologists in the early detection of breast cancer. CAD can automatically indicate suspicious regions in mammographic images, providing a “second opinion” to the radiologist.⁵⁻⁷ Currently, CAD systems are able to detect approximately

---

*Corresponding author: M.A.C.V.: mvieira@sc.usp.br

90% of microcalcifications and masses in digital mammograms. However, in the case of AD detection, the performance is still unsatisfactory, with the detection rate around 50%.^{8,9}

In a previous work, we analyzed several texture descriptors for the detection of Architectural Distortion (AD) in digital mammography using a multi-layer perceptron neural network.¹⁰ In that pipeline, the segmentation step was performed manually, considering AD and normal regions marked by an expert radiologist. However, CAD should perform the whole analysis task in an automatic way to be used in clinical practice. Thus, one of the most important step in this pipeline is the automated analysis of oriented patterns, i.e., the step where the suspicious regions are automatically selected from the original image. Figure 1 shows an usual pipeline of a computer-based system for automated detection of AD in mammography images.¹¹

Figure 1. Simplified version of an usual pipeline for automated detection of architectural distortion in digital mammograms.

Usually, the segmentation step selects about 40 – 100 false-positives (FP) per image. Then, the algorithm performs an unsupervised selection by sorting these regions in descending order, according to the score provided by its mean intensity.^{7,12} The first τ regions of the list are then selected, considering a previously fixed parameter (i.e., a threshold) for further analyzes in the classification step. Thus, regions sorted above this threshold are discarded before going to the classification step.

Consider the mammogram shown at Figure 2a, where the segmentation step inserts a mark, *, into the center of each selected region. In this case, CAD has segmented 71 regions without thresholding. The region highlighted in yellow is the real location of the AD according to the radiologist indication. Setting the threshold to $\tau = 30$, for example, the algorithm will analyze each region, sort and keep only the first 30 ones to send to the classifier. The result after thresholding is presented at Figure 2b. Note that the region containing the AD remains in the list, but there are other 29 false positives (FPs) regions which will be analyzed in the following steps of the CAD pipeline.

Figure 2. Example of mammographic image with automated segmented regions marked by a CAD system (a) without limiting the number of regions in the segmentation step and; (b) limiting the number of regions to $\tau = 30$. The highlighted yellow region shows the real location of the AD in both images.

Limiting the number of regions by a fixed number has the intent to reduce the amount of data to be inserted into the classifier. A high number of FPs regions demands higher computational resources and time. Moreover, the performance of the classification step depends on the number of false positives regions received as input. However, the unsupervised selection of a fixed number of regions (thresholding) could remove the AD region from the list, which will increase the number of false negatives.

Thus, since the initial selection of τ regions depends on the score given to those areas, this paper presents a new method to reduce the number of selected regions, and consequently the FPs rate, based on a candidate ranking given by 11 different measurements. This analysis is performed over the intensity image, Gabor magnitude image and several images in the feature domain (texture map). The number of regions to be selected (threshold value) was increased until the True Positive Rate (TPR) was 100%. It means that all AD regions were correctly selected and the false negative rate was zero.

2. MATERIALS AND METHODS

The usual approach for a computer-based detection of AD in digital mammograms was shown in Figure 1^{7,11,12} and it is shown again in Figure 3, shadowed in blue. In the usual approach, the mammogram is pre-processed to remove the background and the pectoral muscle (for MLO views). Then, the breast image is submitted to a domain transform by applying the Gabor filter in order to enhance the breast structures and perform an analysis of oriented pattern for convergence centers.¹¹ The pre-selected regions are sorted in descending order based on its mean intensity, then the thresholding step is performed according to a fixed τ parameter. The remaining τ regions are then submitted to a trained classifier. More details about the usual approach can be found in its respective paper.^{7,11,12}

Figure 3. Usual (in green) and proposed (in blue) approaches for automated detection of architectural distortion in digital mammograms.

The new approach proposed in this paper, is shown in Figure 3, shadowed in green. We proposed three main modifications in the usual method:

1. The algorithm to analyze oriented patterns use as input the intensity image;
2. Each pre-selected suspicious regions is submitted to a texture feature extractor;
3. The ranking of suspicious regions, in the feature domain, is made using 11 new statistical measurements in addition to the Mean.

The following subsections have the aim to explain in details the novelty of the proposed methodology as well as the dataset used.

2.1 Clinical Image Dataset

The dataset used for the evaluation of the proposed approach is composed by 74 clinical Full-Field Digital Mammography (FFDM) images acquired at the Institute of Radiology of São Paulo (*Instituto de Radiologia de São Paulo - InRAD*) at the University of São Paulo Medical School (*Faculdade de Medicina da Universidade de São Paulo - FMUSP*). This study has IRB approval (number #1.581.220 - *Certificado de Apresentação para Apreciação Ética* - CAAE #56699016.7.0000.0065). Each image, see an example in Figure 4a, has only one region with AD, which were marked by a radiologist.

Figure 4. Example of mammography images used in this work in four stages of processing: (a) intensity image; (b) after breast segmentation and pre-processing; (c) Gabor magnitude response and; (d) “feature map” or “texture image” calculated using the CS-LBP descriptor. Both images (c) and (d) were enhanced for better visualization.

2.2 Texture Image

The schematic presented in Figure 3, shows that the pre-selected regions are used as input to a texture descriptor. In this work we evaluated three texture descriptors: Local Binary Pattern (LBP),¹³ Center-Symmetric Local Binary Pattern (CS-LBP),¹⁴ and the Local Mapped Pattern (LMP).¹⁰

Figure 5. Example of local binary analysis considering (a) a small region of an image with central pixel n_c and its neighborhood, how the (b) LBP and (c) CS-LBP gives a code to the central pixel based on the neighborhood.

The LBP is a well-known local micro-pattern texture descriptor. It is considered one of the most powerful texture descriptor due to its capacity of discrimination, simplicity and low computational cost.¹³ Considering a small portion of an image, as presented in the Figure 5a, and the analysis of a central pixel n_c , the LBP code is calculated based on its neighborhood according to Equation 1.

$$\begin{aligned} \mathbf{LBP} = & s(n_0 - n_c) \times 2^0 + s(n_1 - n_c) \times 2^1 + s(n_2 - n_c) \times 2^2 + \\ & s(n_3 - n_c) \times 2^3 + s(n_4 - n_c) \times 2^4 + s(n_5 - n_c) \times 2^5 + \\ & s(n_6 - n_c) \times 2^6 + s(n_7 - n_c) \times 2^7. \end{aligned} \quad (1)$$

Since the CS-LBP was created upon LBP idea, there is a slight difference between them. It consists in verify the intensity of symmetric pixels in the neighborhood of n_c . Considering the same 3×3 image (see Figure 5a), the CS-LBP code of the central pixel n_c is defined as shown in Equation 2.

$$\begin{aligned} \mathbf{CS-LBP} = & s(n_0 - n_4) \times 2^0 + s(n_1 - n_5) \times 2^1 + \\ & s(n_2 - n_6) \times 2^2 + s(n_3 - n_7) \times 2^3. \end{aligned} \quad (2)$$

The LMP, if compared to LBP, has the big possibility of using any mapping function, i.e., the way how the central pixel will be modified based on its neighborhood.¹⁰ The mathematical formulation of LMP considering a neighborhood W and histogram of B bins is given by the Equation 3.

$$\mathbf{LMP} = \text{round} \left(\frac{\sum_{i=1}^W f_g \cdot P(i)}{\sum_{i=1}^W P(i)} (B - 1) \right), \quad (3)$$

where $P(i)$ is a weight matrix and f_g the mapping function, equations 4 and 5, respectively.

$$P(i) = \begin{bmatrix} 1 & 2 & 4 \\ 128 & 0 & 8 \\ 64 & 32 & 16 \end{bmatrix} \quad (4)$$

$$f_g = \frac{1}{1 + e^{\frac{-(n_i - n_0)}{\beta}}} \quad (5)$$

The configuration of Sigmoid function, f_g in Equation 5, is given by its inclination handled by the parameter β . In previous work was evaluated several configurations of this parameter, however here is reported only the configuration that produced the best results.

All descriptors (LBP, CS-LMP and LMP) creates a code for each pixel of the image to be described by the method. An example of a resulting texture image is presented in the Figure 4d. Our approach consists in apply the descriptor only to the suspicious Regions of Interest (ROIs), as stated previously in the Figure 3. After create a texture map for each ROI, the algorithm applies local statistical measurements in order to score the regions to perform the FPs reduction.

2.3 Local Statistical Analysis

Considering the example showed in Figure 2, there is a large number of regions that will be analyzed by the classifier, but only one has architectural distortion. Thus, to reduce the number of false positives, was defined a unsupervised selection of regions, as illustrated in the Figure 6.

Figure 6. Unsupervised selection of regions based on ranking of statistical measurements.

The regions which represents a center of convergence of oriented patterns are analyzed by 11 statistical measurements considered in this work: IQR (Interquartile Range), MAD (Median Absolute Deviation), Range, Variance, Kurtosis, Skewness, Entropy, Contrast, Correlation, Energy and Homogeneity. The last 5 measurements are Haralick's descriptors and thus calculated over the cooccurrence matrices.¹⁵ The use of Haralick's measurements is justify by the reasonable results presented in the previous study.¹⁰

For completeness of this study, according to Figure 6, the ROIs are extracted from the intensity image, Gabor magnitude or texture map, and submitted to the pre-selected measurement. The dark gray squares represent the FPs while the orange ones the True Positives (TP), i.e. the exact location of the architectural distortion. The result of measurement used will score the regions in order to sort them in descending order. The same process is used by the usual approach as presented in the Figure 3 but, in that scenario, the Mean intensity was used as measurement to score the region.

The sorted list of scored regions, one list per image, needs to be reduced in order to eliminate potential FPs. The former approach use a fixed threshold τ . In this new approach, we suggested to reach a True Positive Rate of 100%, to assure that the number of false negatives is zero, with a minimal False Positives per image. Thus, the 11 measurements were analyzed in images of several domains.

3. RESULTS

To perform a fair evaluation, the proposed approach (see the green shadowed part of Figure 3) was compared to the default one^{7,11,12} (blue shadowed part of Figure 3) using the intensity image as well as the Gabor magnitude, both exemplified in the figures 4b and 4c. Thus, the evaluation at total consists of 12 measurements (the default, i.e., Mean, added to 11 new measurements) applied to 5 images (intensity, Gabor magnitude, LBP, CS-LBP and LMP texture maps). For mater of comparison, the threshold was changed in step of 5 until at least one measurement reached 100% of TPR.

Quantization of the results was in terms of True Positive Rate (TPR) of each measure considering the whole dataset of 74 images. The TPR was calculated according to Equation 6:

$$TPR = \frac{\text{number of AD regions correctly selected}}{\text{total number of images (AD)}} * 100 \quad (6)$$

Preliminary results, related to usual approach, are presented in plots of Figure 7, where the Mean measurement applied to Gabor images has reached 100% of TPR with threshold of 85, see Figure 7a. In the same image, if considered other measurements as shown in Figure 7b, it is possible to have equal result with a lower threshold. These results were hit by Range and Variance, with a threshold of 60 regions, which in our dataset corresponds to the average of 40.7 FPs per image.

Figure 7. Results in terms of True Positive Rate based on analysis of regions using 12 measurements over the Gabor image: (a) default approach using the Mean intensity measurement; (b) other 11 proposed measurements.

The proposed approach uses the intensity image and texture maps to perform the FP reduction. The results considering these new images are shown in Figure 8. The case where the analysis were performed over the intensity image is shown at Figure 8a. It suggests that the Contrast measurement is the best parameter to reduce the number of FP per image, as the threshold for a TPR of 100% was 55 ROIs, corresponding to an average number of 36.8 selected regions per image. The same threshold value was enough to the Skewness measurement, which selected all AD regions using the texture image produced by the LBP method.

The Figure 8c shows that Contrast was the best measurement to select all AD regions when the CS-LBP was used with an average of 37.8 ROIs per image (threshold = 60). The LMP with $\beta = 0.2$ has the best result over all other methods. Here the contrast was the best measurement and achieved 100% of TPR with the lowest threshold value: 50 with average of 35.5 ROIs per image.

The best measurement for each type of image were summarized in the Table 1. The shadowed row corresponds to the default approach which reached 100% of TPR with the average of 44.4 (threshold = 85). Correlation measured over the LMP texture image provided the best result in terms of number of average regions per image (35.5).

Table 1. Each approach and its respective setup when the true positive rate achieved 100%. The default approach is highlighted by a shadowed row. The best result in terms of number of ROIs per image is highlighted in bold.

<i>Image/Texture map</i>	<i>Measurement</i>	<i>Threshold</i>	<i>Susp. spots per Image</i>
Gabor mag. ^{7,11,12}	Mean	85	44.4
Gabor mag.	Range/Variance	60	41.7
Intensity	Contrast	55	36.8
LBP ¹³	Skewness	55	36.8
CS-LBP ¹⁴	Contrast	60	37.8
LMP ($\beta = 0.2$)¹⁰	Correlation	50	35.5

In this study we verified that the number of selected ROIs change significantly according to the image. Thus, this information must be considered when deciding which method will be used to select the ROIs. The

Figure 8. Results in terms of True Positive Rate based on analysis of regions using 12 measurements over the images (a) Intensity image; and the texture maps (b) LBP, (c) CS-LBP and (d) LMP.

histograms of Figure 9 show the number of ROIs recovered by the oriented gradient analysis algorithm over the Gabor magnitude image, (Figure 9a) and over the Intensity image (Figure 9b).

Figure 9. Histograms showing the number of suspicious regions automatically selected for each one of the 74 images used in this work after the analysis of oriented gradients over (a) Gabor magnitude and (b) Intensity image.

It is possible to verify that the use of Gabor magnitude selected more ROIs than the Intensity image. However, this difference did not affect the reduction of FPs. The threshold values and the average number of ROIs per image that were selected are shown at Figure 9. According to Figure 9a, the default approach reached 100%

of TPR with a threshold of 85, and almost included all selected ROIs except for two images. Using Range or Variance as the statistical measurement reduced the average number of selected regions without increasing the number of false negatives (TPR = 100% for both cases).

The histogram of Figure 9b shows the number of ROIs selected for each image in the database considering the intensity image instead of the Gabor image. The use of texture images instead of Gabor reduced the number of false positives per image. The best descriptor was the LMP, which selected the average number of 35.5 ROIs per image.

4. CONCLUSIONS

This paper proposed a new method to select candidates of architectural distortion regions in a CAD scheme by reducing the number of false-positives using texture maps and statistical measurements. Reducing the number of FP per image is an essential step in any computer vision pipeline, since the FPs can mask the desired result. Since the AD is well characterized by its textural pattern, we considered three texture descriptors (LBP, CS-LBP and LMP) before sorting the selected ROIs, where each pixel corresponds to a local binary code created after analysis of neighborhood of a central pixel.

Another novelty presented in this study is the use of other statistical measurements instead of the Mean intensity. The usual approach uses the Mean intensity to score and sort suspicious regions before discard the considered FPs. However, other measurement can produce better results, mainly when it is calculated over texture maps.

ACKNOWLEDGMENTS

This project is supported by São Paulo Research Foundation (FAPESP) grant #2015/20812–5 and the National Council for Scientific and Technological Development (CNPq).

REFERENCES

- [1] Banik, S., Rangayyan, R. M., and Desautels, J. E. L., “Detection of Architectural Distortion in Prior Mammograms,” *IEEE Transactions on Medical Imaging* **30**(2), 279–294 (2011).
- [2] Ray, K. M., Turner, E., Sickles, E. A., and Joe, B. N., “Suspicious Findings at Digital Breast Tomosynthesis Occult to Conventional Digital Mammography: Imaging Features and Pathology Findings,” *The Breast Journal*, 1–5 (2015).
- [3] Bahl, M., Baker, J. A., Kinsey, E. N., and Ghatge, S. V., “Architectural Distortion on Mammography: Correlation With Pathologic Outcomes and Predictors of Malignancy,” *American Journal of Roentgenology* **205**(6), 1339–1345 (2015).
- [4] Gaur, S., Dialani, V., Slanetz, P. J., and Eisenberg, R. L., “Architectural Distortion of the Breast,” *American Journal of Roentgenology* **201**, 662–670 (2013).
- [5] Dean, J. C. and Ilvento, C. C., “Improved Cancer Detection Using Computer-Aided Detection with Diagnostic and Screening Mammography - prospective study of 104 cancers,” *American Journal of Roentgenology* **187**, 20–28 (2006).
- [6] Yang, S. K., Moon, W. K., Cho, N., Park, J. S., Cha, J. H., Kim, S. M., Kim, S. J., and Im, J.-G., “Screening mammography-detected cancers: Sensitivity of a computer-aided detection system applied to full-field digital mammograms,” *Radiology* **244**(1), 104–111 (2007).
- [7] Casti, P., Mencattini, A., Salmeri, M., Ancona, A., Mangeri, F., Pepe, M., and Rangayyan, R., “Contour-independent detection and classification of mammographic lesions,” *Biomedical Signal Processing and Control* **25**, 165–177 (2016).
- [8] Baker, J. A., Rosen, E. L., Lo, J. Y., Gimenez, E. I., Walsh, R., and Soo, M. S., “Computer-aided detection (CAD) in screening mammography: Sensitivity of commercial CAD systems for detecting architectural distortion,” *American Journal of Roentgenology* **181**(4), 1083–1088 (2003).
- [9] Rangayyan, R. M., Banik, S., and Desautels, J. E. L., “Computer-Aided Detection of Architectural Distortion in Prior Mammograms of Interval Cancer,” *Journal of Digital Imaging* **23**(5), 169 (2010).

- [10] Oliveira, H. C. R., Moraes, D. R., Reche, G. A., Borges, L. R., Catani, J. H., de Barros, N., Melo, C. F. E., Gonzaga, A., and Vieira, M. A. C., "A new texture descriptor based on local micro-pattern for detection of architectural distortion in mammographic images," in [*Proc. of SPIE*], **10134**, 101342U–101342U–9 (2017).
- [11] Banik, S., Rangayyan, R. M., and Desautels, J. L., [*Computer-aided Detection of Architectural Distortion in Prior Mammograms of Interval Cancer*], vol. 8, Morgan & Claypool Publishers (2013).
- [12] Banik, S., Rangayyan, R. M., and Desautels, J. E. L., "Detection of Architectural Distortion in Prior Mammograms," *IEEE Transactions on Medical Imaging* **30**(2), 279–294 (2011).
- [13] Ojala, T., Pietikainen, M., and Maenpaa, T., "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(7), 971–987 (2002).
- [14] Heikkilä, M., Pietikäinen, M., and Schmid, C., "Description of interest regions with local binary patterns," *Pattern Recognition* **42**(3), 425–436 (2009).
- [15] Haralick, R. M., "Statistical and structural approaches to texture," in [*Proceedings of the IEEE*], **67**(5), 786–804 (1979).

7.4 Artigo 4: *Biomedical Signal Processing and Control*

Artigo publicado na revista

Biomedical Signal Processing and Control.

<https://doi.org/10.1016/j.bspc.2019.01.001>

A cross-cutting approach for tracking architectural distortion loci on digital breast tomosynthesis slices

Helder C.R. de Oliveira^a, Arianna Mencattini^{b,*}, Paola Casti^b, Juliana H. Catani^c,
Nestor de Barros^c, Adilson Gonzaga^a, Eugenio Martinelli^{b,**}, Marcelo A. da Costa Vieira^a

^a University of São Paulo - São Carlos School of Engineering, São Carlos-SP, Brazil

^b University of Rome - Tor Vergata, Rome, Italy

^c University of São Paulo - Medical School, São Paulo, Brazil

ARTICLE INFO

Article history:

Received 12 July 2018

Received in revised form 1 November 2018

Accepted 4 January 2019

Keywords:

Architectural distortion

Digital breast tomosynthesis

Breast cancer

Computer aided detection

Gabor filter

Cell tracking

ABSTRACT

Background and objective: Full-field digital mammography (FFDM) is the predominant breast cancer screening exam used. However, with the emergence of digital breast tomosynthesis (DBT) the radiologists could improve early recognition of breast cancer signs. In this scenario, the detection of architectural distortion (AD) is still a challenging task. ADs are very subtle contraction of the breast parenchyma that could represent the earliest manifestation of cancer, assessing at present 50% of missed cases. **Methods:** This paper proposes a new paradigm to detect AD in DBT exams by a cross-cutting approach exploiting the 3-dimensionality of the imaging modality. After locating AD candidates in each DBT slice, the suspicious spots are tracked in cross-slice direction and then characterized in terms of neighboring texture. In this approach, which mimics radiologist's scrolling down over zoomed slices, we reduce the amount of uninformative signs collected in DBT exams by preserving the large variability of AD appearance. **Results:** Using 37 sets of DBT slices containing at least one AD locus indicated by a radiologist, the proposed methodology reaches an AUC of 0.84, with only one false negative exam at sensitivity of 0.9. **Conclusions:** The results show that the proposed algorithm can be a promising tool for the automatic detection of AD loci. Future work will address the extension of the dataset of DBT slices as well the improvement of algorithm performance toward the application in the clinical practice.

© 2019 Published by Elsevier Ltd.

1. Introduction

Breast cancer is the main cause of cancer death among women worldwide with nearly 1.7 million new cases reported by the World Health Organization (OMS) [1]. The probability of successfully treating the disease increases up to 30% when detected at the initial stage, whereas treatment at an advanced stage is less effective and prone to severe human physical discomfort [2–4]. Early detection can be achieved with screening programs by means of full-field digital mammography (FFDM) [5–7].

Unfortunately, FFDM may provide inaccurate results, mainly for high dense breasts, for which misinterpretation can occur due to overlapping-tissues effects, where the 2D projection of the fibroglandular tissue can hide a lesion (false negative) or resemble non-existing anomalies (false positive) [8–10].

Because of the advent of Digital Breast Tomosynthesis (DBT) and its clinical use since 2011 [8,11,12], the mentioned limitations have been partially overcome. DBT is an imaging modality of the breast where the images (projections) are acquired at specific projection angles while the X-ray tube moves around the breast over a limited angular range. Based on these projections, a computational system of tomographic reconstruction is used to generate a three-dimensional volume of the breast and to produce a set of tomographic slices [8,9,13,11] that can be presented to the radiologist for analysis and reporting. Moreover, the total radiation dose is set to be comparable to that used in FFDM.

Recent studies have shown that DBT not only minimizes the problem of tissue overlapping, but it also ensures increased sensitivity and specificity to abnormalities if compared to FFDM [14–17]. Therefore, it will offer a promising alternative to FFDM-based screening tools in the near future [14–16].

Architectural distortion (AD) is a subtle contraction of the breast parenchyma that can indicate a lesion at a very early stage of formation, in most cases, appearing before the formation of any mass or calcification [18–22]. Some studies have shown that if the AD is

* Corresponding author.

** Co-last authorship

E-mail address: mencattini@ing.uniroma2.it (A. Mencattini).

Fig. 1. A sketch of the novel proposed paradigm.

recognized, the probability of a correct prognosis of cancer treatment is higher if compared to early detection of masses and/or calcifications [18,19].

ADs are classified as BI-RADS 4C [23] since they can represent the onset of malignant cancer later verified by biopsy. ADs are very hard to detect in mammographic images because of roughly defined border and no visible masses [15]. In most cases, the subtlety of such parenchyma phenotyping impairs AD detection, even to experienced radiologists [15], reporting a human average detection rate of AD nearly 50%, which makes AD the most common finding in retrospective exams of false-negative cases [18,19].

In order to aid radiologists in the early diagnosis of breast cancer, functionalities of Computer-Aided Detection (CAD) systems are going to be even more exploited [24–26]. Several studies have shown that radiologists, when assisted by CAD, are capable of detecting about the 90% of macrocalcifications and masses. However, performance in automatic detection of AD, even with the use of DBT is still unsatisfactory, with detection rates just around 60% [27,28,17,7,29,30]. Anyhow, a few works addressed the topic of detecting architectural distortion alone (i.e., most CAD systems combine the detection of spiculated masses and architectural distortion), thus suggesting the fact that a dedicated method may improve performances.

Thus, the aim of this paper is to propose a new paradigm to detect AD in DBT exams by a cross-cutting approach exploiting the 3-dimensionality of the imaging modality.

1.1. Main contribution

A new paradigm (see Fig. 1) is presented here in order to extract relevant information from DBT volume. First, we localize AD candidates in each DBT slice (blue stars over each slice) by exploiting a modified procedure based on the combination of Gaussian curvature computation [26] and Gabor filtering approaches.

Second, in order to reduce the increased amount of data to process, we configure an *AD path* along the breast volume, by computing an AD trajectory from the automatically detected coordinates located at each slice (the red path across the slices connecting the candidates). Free-available Fiji/ImageJ [31] software, suitable tuned, was applied for tracking.

Third, each candidate trajectory is then characterized using statistical and textural descriptors computed along and in the neighbourhood of the track. The derived descriptors are then used to train a supervised AD-path recognition model. This new scenario

emulates radiologists in scrolling down the cursor over different zoomed DBT slices of the same breast.

As a primary outcome we provide candidate locii for AD within each slice, and further we construct the AD candidate's path along the breast tissue.

The proposed new way of reasoning allows the radiologist to orientate himself in the labyrinth of information contained in the DBT slices by means of a cross-cutting approach. Particles tracking experiences, widely exploited in the field of cell cross-talk and drug delivery [32,33], are here applied to locate first, and follow then, the spatial developing of the parenchymal distortion of the breast tissue. Such approach allowed us to improve the performance of the CAD system in recognizing AD locii in absence of dense kernels (i.e. massive lesions).

1.2. Related work

Table 1 lists the most relevant work on the detection of AD in breast images. As it can be seen, there is a high heterogeneity in the methods proposed for the scope. Almost all of them use 2D mammography (FFDM and screen-film mammographies). Besides that, some approaches concern the training of a classifier in discriminating normal ROIs from those ones containing an AD. ROIs are extracted manually based on previous assessment made by radiologists. Available information regarding the number of images and of ROIs, the number of false positives per image, sensitivity, and area under the ROC (AUC) curve is listed.

The state-of-the-art work listed demonstrate the difficulty of the AD detection task. In fact, most of the work addressed the simplified problem of discriminating ROIs, previously extracted, and not the much more critical task of localizing the AD locii in the breast. Attempts by Rangayyan et al. [37], Casti et al. [26], and Nemoto et al. [34] try to directly solve the problem of finding AD candidates without human intervention. The first actual attempts with DBT have been addressed in [38], by traducing the problem of detecting AD locii in DBT in the similar problem of finding stellate massive lesions. In their approach, pathological tissue containing a general abnormality (e.g., mass or AD) is classified using an *a contrario* approach able to firmly recognize normal parenchyma in each slice. Aggregation at feature level is then achieved in order to provide a volumetric information about the abnormality. Unfortunately, despite the elegance of the approach, the number of ADs used for validating the strategy is still quite small (3 ADs); therefore it is difficult to fully verify the capability to find AD locii not showing

Table 1
Relevant works addressing the detection of ADs in the breast region. Results as summarized in terms of False Positives (per Image, FPpl, or per Patient, FPpP), sensitivity, and area under the ROC (AUC) curve, when available. DDSM: Digital Database for Screening Mammography; MIAS: Mammographic Image Analysis Society database.

Author	Year	Dataset	FPs	Sensitivity	AUC
Nemoto et al. [34]	2009	25 digitized mammograms with AD	0.8 FPpl	80%	NA
Jasionowska and Przelaskowski [35]	2010	292 DDSM ROIs – 34 with AD – 258 w/o AD	0.86 FPpl	68%	NA
Minavathi et al. [36]	2011	291 ROIs from DDSM and MIAS – 42 with AD – 249 w/o AD	NA	94%	0.90
Rangayyan et al. [37]	2013	158 digitized mammograms (106 prior mammograms of interval-cancer cases and 52 normal mammograms)	3.4 FPpP	80%	NA
Palma et al. [38]	2014	101 DBT cases (53 with cancer – 3 AD – and 48 w/o)	1.31	80%	NA
Kamra et al. [39]	2015	348 ROIs from DDSM and MIAS – 165 with AD – 183 w/o AD	NA	93%	0.95
Casti et al. [26]	2016	116 mammograms from DDSM (with 98 malignant ADs and 24 benign ADs)	5.41 FPpl	81%	NA
Liu et al. [40]	2017	138 ROIs from DDSM (69 with AD and 69 w/o AD)	NA	92%	NA
Oliveira et al. [41]	2017	400 FFDM ROIs (200 with AD and 200 w/o AD)	NA	NA	0.97
Bahl et al. [7]	2017	395 DBT cases (274 with ADs)	NA	65.7 %	NA
Alshafey et al. [17]	2018	121 DBT cases (121 with ADs)	NA	48.8 %	NA

a dense kernel. Recent retrospective studies tried to assess superiority of DBT over digital mammography in ADs finding. Expert radiologists have been asked to locate ADs retrospectively and results were compared with biopsy-guided results. Sensitivity values settle on [50%–70%] that constitute a reference point for the present investigation.

2. Materials and methods

2.1. Digital breast tomosynthesis dataset

The dataset used is composed of 20 DBT exams (37 volume sets) obtained from the Institute of Radiology of São Paulo (*Instituto de Radiologia de São Paulo – InRAD*) at the University of São Paulo Medical School (*Faculdade de Medicina da Universidade de São Paulo – FMUSP*), under Institutional Review Board (IRB) approval.¹ Since in a DBT exam the radiologist normally analyzes the images corresponding to the slices, which were produced by the reconstruction algorithm over the projections, only this category of images was considered in this work. All set of images were acquired using the Hologic Selenia Dimensions^{TM} Mammographic System (see Fig. 1). Each exam usually comprises a set of *craniocaudal* (CC) and/or *mediolateral oblique* (MLO) views of both breasts. As a preliminary study, in this work we considered 37 sets of DBT slices containing at least one AD locus, manually indicated within a circle drawn by expert radiologists in each DBT slice. The number of images (slices) of each set can vary according to the compressed breast thickness. Each slice has a reconstructed pixel size of approximately 100 μm and corresponds to a cross section of the breast, 1 mm thick, acquired parallel to the breast support plate. Histological examinations of the lesions were not available at the time of the acquisition, but future work will be devoted to the further investigation of automatic discrimination between malignant and benign ADs [17,7].

As a visual demonstration of the complexity of the task, Fig. 2 compares two sets of 12 ROI-stack taken at distance of three slices (about 3 mm), cropped around a reported AD locus (upper part of Fig. 2) and in normal breast tissue (lower part of Fig. 2). Note that, in the case shown, the presence of the AD locus is quite visible thanks to the spicules radiating from a central point and, most importantly, thanks to the fact that this pattern tends to disappear as long as slices move away from the initial slices. Such behaviour does not occur in presence of normal tissue, where, instead, the characteristic patterns remain visible through the volume.

3. The proposed approach

Fig. 3 summarizes the processing steps used in this work.

3.1. Pre-processing

In the first step, a 3D stack of all the slices of the same set is built following the cut order. A contrast enhanced procedure based on the linear transform encoded in the Look-Up Table (LUT) is then applied to each image of the stack.

Due to the anisotropy in the construction of the image stack (i.e., image spatial resolution is different from slices distance), a procedure to assure isotropy of the breast volume acquired is a necessary next step [42]. This is due since the necessity of perform tracking across slices. The whole volume was resized in every direction to 300 μm using bilinear interpolation [42,43]. Such step has the objective to reduce finer details that could represent noise as well

¹ IRB approval #1.581.220 - Certificado de Apresentação para Apreciação Ética – CAAE #56699016.7.0000.0065.

Fig. 2. Two examples of ROI-stacks extracted around an AD locus (upper part) and in normal tissue (lower part). Slices at a distance of about 3 mm each were collected.

as to keep useful information which could be lost by coarser pixel [44]. Isotropy allows us to process the volume in any direction without any further rescaling. Fig. 4 illustrates the steps of the main DBT processing to the scope of AD candidate localization.

3.2. Breast segmentation

Each DBT slice contains a considerable portion of background, no matter the size of breast imaged. The background itself corresponds to the area beyond the breast border, and therefore, due to its uninformative content, a background removal procedure is needed to avoid useless processing.

On the other hand, some portions of the breast, such as the pectoral muscle, can be removed from MLO images with the aim to reduce confounding effects during processing.

We apply here the algorithms proposed in [45] and exhaustively assessed in a most recent work [26]. Fig. 4(a) presents the result of the segmentation algorithm (the breast portion is highlighted by the green border).

3.3. Detection of suspicious spots

First, Gabor filtering [46] has been applied in order to highlight oriented patterns in the breast tissue, Fig. 4(b). A set of 18 multidirectional Gabor filters was used to extract various directional components of the breast region. Consider the real Gabor kernel, $g(x, y)$, oriented at $-\pi/2$, defined as

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2}\right)\right] \cos(2\pi xf) \quad (1)$$

where σ_x and σ_y are the standard deviation values along the x and y directions, respectively, and f is the spatial frequency, inversely linked to the thickness parameter τ and to the σ_x , as follows $\tau = 1/f = 2\sigma_x\sqrt{2\log(2)}$. Multidirectional analysis can be obtained by simply rotating the Gabor kernel in Eq. (1) over the angular range $(-\pi/2, \pi/2]$. A maximum response approach is then applied to the multidirectional-filter output to recover a unique magnitude and a phase response at each pixel [47], as shown in Fig. 4(b). The values of σ_x and σ_y are reciprocally related to the elongation parameter

$l = \sigma_y/\sigma_x$, which was set to 3. The thickness parameter τ was instead set to 24 pixels (7.2 mm) leading to a value of $\sigma_x = 10$ pixels (3 mm). Correspondingly, we have $\sigma_y = 30$ pixels (9 mm).

As a second step, we modified an approach already validated on mass detection in mammography [26]. The main idea of AD locii detection is to locate the center of local intensity convexities, while reducing coarser information. To perform this, a Gaussian filter with $\sigma = 12$ pixels (2.4 mm) was applied to reduce noise and retain only low-frequency information related to the cores of the AD. Formally, the intensity values of a given image can be characterized in terms of a Monge patch, $\eta : D \subset \mathcal{R}^2 \rightarrow \mathcal{R}^3$, defined as:

$$\eta(x, y) = (x, y, h(x, y))$$

where D is the domain of the two space variables, x and y , and $h(x, y) : D \rightarrow [0, 1]$ is a function that associates to each spatial coordinate (x, y) an intensity value h in $[0, 1]$ [48]. The Gaussian curvature, K , defined as the product of the two principal curvatures of a surface in any point, determines the topology of any given surface. Let $\mathbf{p} = (x_0, y_0)$ be the coordinates of a generic point on η then, a topology classification can be performed as follows

$$\begin{aligned} \text{if } K(x_0, y_0) &> 0, & \text{then } \mathbf{p} \text{ is elliptic,} \\ \text{if } K(x_0, y_0) &< 0, & \text{then } \mathbf{p} \text{ is hyperbolic,} \\ \text{if } K(x_0, y_0) &= 0, & \text{then } \mathbf{p} \text{ is parabolic or planar,} \end{aligned}$$

Thanks to the assumptions on the range of variation of h (i.e., $[0, 1]$, that bounds the variation on its partial derivatives) and on the smoothing effects of the preliminary Gaussian filtering [26], then K can be expressed as

$$K = \det[\text{Hess}(h)],$$

where $\text{Hess}(h)$ is the Hessian operator applied to h . If an orthonormal basis of eigenvectors for $\text{Hess}(h)$ is chosen, K can also be expressed via the product of the corresponding eigenvalues, λ_1 and λ_2 . The main assumption here is that the AD locii $(\tilde{x}, \tilde{y})$ satisfy the condition $\lambda_{1,2}(\tilde{x}, \tilde{y}) < 0$. These condition leads to the selection of a set of focal areas on the slice that corresponds to the nodes of a convergent gradient vector field (GVF), as shown in Fig. 4(c). For

Fig. 3. Workflow illustrating a synthesized version of the steps used in our approach.

Fig. 4. Examples of slice by slice processing steps: (a) breast (contour and pectoral muscle) automatic segmentation; (b) Gabor filtering to enhance directional patterns; (c) detection of suspicious spots (contoured regions). The white spot indicates the center of the region.

Fig. 5. A sketch of the AD tracking and ROI-processing algorithm. (a) A set of 18 tracks moving across the slices are extracted using TrackMate. The red track represents the true AD path, the remaining tracks are all false paths. (b) For each track, a region of size 100×100 pixels is extracted around each candidate point of the track, thus leading to a collection of ROIs (ROI 1 to ROI 150 in the figure). (c) Finally, a feature extraction procedure is applied separately to each ROI and then standard statistical aggregation operators are applied to extract a synthetic track-associated descriptor. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

each region, the centroid's coordinates within each convex regions were taken as the representative candidates for the AD focus.

The main novelty of the proposed localization approach is the combination of pipeline of Gabor filtering and Gaussian curvature approach, that extends to spiculated findings the mass localization approach [26].

3.4. Tracking of suspicious spots

During DBT visual evaluation, radiologists scroll the set of images of the exam in cross-slice direction in order to see how the breast parenchyma changes at different deepness of the breast. Based on this well established protocol, we intend to develop a computational system able to mimic such a strategy.

Assuming that the focal spots are part of an AD lesion, it is expected that the suspicious area continues in a sequence of neighboring slices. In order to track all pointed areas in cross-slice direction, a particle tracking software tool was used. TrackMate is an open-source software developed to track single particles [49] shared as plugin of popular Fiji/ImageJ [31] tool, an image editor widely used by researchers working on biomedical images.

Based on all the spots localized in each slice, TrackMate is capable of performing a preliminary tracking of the AD path along different slices.

The TrackMate setting used in the present work is described below. TrackMate parameters were chosen heuristically, as those providing reasonable tracking results. The parameters changed from the default setting are reported in Table 2.

- are based on Difference of Gaussian (DoG) algorithm, where the second Gaussian filter has a largest standard deviation if compared to the first one. Subtracting one resulting image from another is possible to find local maxima which will be considered as the particle locations;
- after detection of potential particle sites, those “false ones” are removed based on features like min, max, mean, median, total intensity, standard deviation, contrast to noise ratio, signal to noise ratio or estimated diameter. Based on the thresholding performed on the selected features, the user can discard candidates;
- the remaining spots are tracked based on the Linear Assignment Problem (LAP) algorithm proposed in [50]. The default tracker, the *Simple LAP Tracker*, is suitable to situations where the particles

Table 2
TrackMate parameters.

	Parameter	Value
<i>Detector</i>	Subpixel localization	True
	Median filtering	False
	Threshold	10
	Radius	5
	Target channel	1
<i>Tracker</i>	Algorithm	Simple LAP
	Linking max. distance	20
	Allow gap closing	True
	Gap closing max. distance	10
	Max. frame gap	10
	Allow track splitting	False
	Splitting max. distance	15
	Allow track merging	False
	Merging max. distance	15

do not merge or split and is recommended to situations where a Brownian motion exists. At the end of this process the tracks are built;

The tracking result consists in a list of track points saved in a XML file for further analysis. It is worth to mention that, all the settings defined can also be saved in another XML file.

Fig. 5(a) presents some examples of tracks in a 3D plot. A set of 18 coloured tracks, extracted using TrackMate, moving across four slices of example are shown. The red track represents the true AD path while the remaining tracks are all false AD paths.

3.5. Feature extraction

In order to discriminate AD tracks from normal tissue, we extracted some texture features from the region surrounding the track. Consider the red track in Fig. 5(b). First, a rectangular region of size 100×100 pixels (3×3 cm) is extracted around each point of the candidate track in each slice (i.e., the shaded pink blocks around the track); second, Fig. 5(c), standard statistical descriptors are computed from the intensity values and from the Gabor magnitude values in this rectangular region. The two domains allow to jointly describe peculiar characteristics of AD and normal tracks.

Since a single descriptor value is extracted in each slice, at the end we collected a set of values for each track. A standard way to

Fig. 6. Three different examples of intensity and Gabor magnitude ROI stacks for diversified f_6 values and pattern behaviour. From top to bottom: (a) kurtosis values over the ROI stack 1 in (b) (selected as having the f_6 values lower than the 30th percentile over the same slices-set) and visually represented by almost homogeneous dense AD pattern; (c) kurtosis values corresponding to the ROI stack 2 in (d) that contains normal tissue having an average f_6 value in [30–70]th percentile over the same slices-set; (e) kurtosis values computed over the ROI stack 3 in (f) that correspond to f_6 values larger than 70th percentile over the same slices-set, that visually correspond to strong radial patterns and/or spicule track positioning. (g) Boxplots of the values of f_6 in the three ROI stacks. The p-values of the ANOVA test with Bonferroni's correction for multicomparison experiment are also indicated along with the average f_6 values (number in red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

aggregate such values, thus producing *aggregated descriptors*, was to compute the average and the standard deviation of the collected *slice-based descriptors* over the track.

From preliminary data analysis, we selected the following parameters: the average of the interquartile range (f_1), of the average values (f_2), of the kurtosis values (f_3), and of the entropy values (f_4) computed using the image intensity.

In addition, we included the standard deviation of the entropy values computed using the image intensity (f_5) and the average kurtosis values computed using the Gabor magnitude (f_6).

3.6. False AD track reduction

One of the six descriptors, f_6 , i.e., the average kurtosis computed over the Gabor magnitude of the ROI stack, was used to perform preliminary unsupervised false tracks reduction.

Considering Fig. 6, three facts were observed: first, very low values of f_6 indicated, on average, a collection of ROIs with flatter Gabor magnitude values distribution (low kurtosis). This happened when the track passed close to the centers of an almost dense AD locus, exhibiting a local intensity homogeneity (case ROI stack 1 in Fig. 6). On the other hand, very high values of f_6 indicated a stronger tailedness of the distribution of Gabor magnitude values that occurred when the track passes on a spicule of the AD in presence of a spiculated AD pattern (case ROI stack 3 in Fig. 6). The remaining values of f_6 (medium intensity) were likely to belong to an average heterogeneity typical of normal tissue (case ROI stack 2 in Fig. 6). Note that for the sake of compactness, only five ROI stacks are illustrated for each track. The distribution of the kurtosis values of all the stacks of the three tracks have been also compared using a multicomparison ANOVA test with Bonferroni's correction. The corresponding p -values are also indicated along with the average f_6 average values (numbers in red). In summary, the tracks with very low or very high f_6 values were kept as possible candidates of AD pattern. The reduction process is slices-set-dependent (i.e., the percentile of f_6 is computed over the tracks of a given set of slices), hence it is not possible to eliminate all the tracks in a given image at this step.

The approach could also allow to locate dense kernel patterns (low average kurtosis values) that correspond to massive lesions that can appear in conjunction with an AD finding.

3.7. Machine learning for AD tracks recognition

In order to automatically recognize true ADs tracks from normal ones, we trained a Quadratic Discriminant Analysis (QDA) model on the six features computed on an independent partition of the exams in the dataset and test it to the remaining portion.

4. Experimental set-up

A set of 20 clinical exams have been considered for validating the whole method. The exams correspond to 37 sets of DBT slices containing at least one AD locus, for a total of 711 tracks extracted, of which 91 are true tracks. We considered as true a track falling within a circle manually drawn by expert radiologists, leading to a total number of tracks about three times larger than the expected number of ADs. Anyway, this is compliant with the proposed approach since, in presence of highly spiculated ADs, tracks may describe different portions of the spicules and be more than the number of ADs loci. Hence, the initial number of false positive tracks per breast volume (FPtpBV) is 16.75 with a number of false positive tracks per exam equal to 31. Some exams include the cranio-caudal (CC) and the mediolateral-oblique (MLO) views if the radiologist reported the AD loci in both views. The low number of images available forced us to use a subset of these images for feature selection and tracking parameters setting. Training and testing

Fig. 7. Free-response ROC (FROC) curve. Operating points at 0.9, 0.8, and 0.7 sensitivity were highlighted by the red lines along with the corresponding FPtpBV values. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of the QDA model have been then performed on the same dataset of images. To do this, a leave-one-exam-out (LOEO) cross-validation procedure was used to test the method by an exam-independent procedure. Results were quantified in terms of Area Under the ROC (AUC) curve and the Free-response ROC (FROC) curve. The former indicator provides a general relationship of sensitivity versus specificity on the entire dataset. The latter representation provided us a measure of the method's sensitivity vs. the number of FPtpBV resulted, thus quantifying the number of expected false tracks residual in each image. Also, we provide a quantification of the number of False Negative exams (FNe) to evaluate the robustness of the method to that annoying problem.

Simulations have been performed on a quad-core Intel® core™ processor i7-7500U CPU 2.70 GHz, with 16 GB of RAM. Average (standard deviation) time consuming required to process a single slice was 14.46 (3.4) s.

5. Results

In this section, we described the whole experiment run to verify the effectiveness of the proposed approach in recognizing true AD tracks in the DBT slices. First, an unsupervised false AD track reduction procedure has been applied on each DBT slice, Section 3.6, and then, a supervised machine learning procedure was used in order to recognize tracks identifying an AD, Section 5.1.

After the application of the unsupervised AD tracks reduction, tracks were reduced from 711 to 428 with residual 67 true tracks. Hence, the number of FPtpBV was reduced from 16.7 to 9.76 with no further loss of positive exams.

The five remaining descriptors are then used in a machine learning task for the supervised discrimination of true AD tracks from normal breast tissue distortion.

5.1. Supervised AD tracks recognition results

The method was successfully validated on the 20 exams producing the FROC curve graphed in Fig. 7. Correspondingly, the obtained AUC value was equal to 0.84.

Regarding the number of FNe for the different levels of sensitivity, we obtain the results reported in Table 3.

Fig. 8. Some visual examples of AD tracks recognition process: red tracks represent true positive tracks, green tracks identify false positive tracks, cyan tracks represent false negative tracks, and blue tracks identify true positive tracks. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Result of supervised recognition of tracks in terms of sensitivity and number of False Negative exams (FNe).

Sensitivity	0.7	0.8	0.9
FNe	3	2	1

It is important to note that at a sensitivity level of 0.8 we only lost two exams while providing a number of false tracks per breast volume slightly greater than 2. Remarkable is the fact that the supervised AD tracks recognition reduces the number of FPtpBV from an initial value of 16.7 to an average final value of 2.35 at the expense of only two lost exams.

In order to provide some visual examples of the results of AD tracks reduction, in Fig. 8 we illustrate some examples of the procedure by graphing tracks with different colors according to the final assessment provided by the system. More specifically, red tracks identify true positive tracks, green tracks represent false positive tracks, cyan tracks identify false negative tracks, while blue tracks locate true negative tracks.

Note that in some cases (example on the left in Fig. 8) the false negative track (cyan track) does not influence substantially the reliability of the approach. In fact, the lost track (in cyan) is actually a peripheral part of the same AD that has been correctly located (red track). In other situations, as in the right part of Fig. 8, more than one track represents the same AD. In this case in fact, the AD region is less localized and more extensive than the previous case, thus motivating the localization of more than one track for the same AD. In such cases, performance were computed by collapsing such multiplicity in only one candidate.

6. Discussion and future directions

6.1. Comparative results

In this section, we present comparative results obtained by using alternative image transform and feature extraction strategies. In light of the emerging trends in Deep Learning (DL) architectures [51], we compared the proposed approach with a ROI-based processing approach. The ROIs extracted around the tracks and used in the presented method to compute features for pattern classification were directly fed to the pre-trained DL network, Alexnet [52]. Alexnet was then used for transfer learning processing in order to transform each ROI into a matrix of numerical descriptors from

Table 4
Comparative results of the proposed approach (third row) with a DL architecture (first row) and LBP strategy (second row). Columns indicate AUC and FPtpBV values at increased sensitivities of 0.7, 0.8, and 0.9.

Method	AUC	FPtpBV		
		sens=0.7	sens = 0.8	sens =0.9
Alexnet on ROIs and QDA	0.78	4.69	5.50	6.69
LBP on ROIs and QDA	0.74	7.52	8.00	8.62
Proposed approach	0.84	1.46	2.35	3.38

pooling layer 2 [53]. In analogy to the proposed methodology, and with the aim to confine the purview of comparison, we used the same supervised machine learning approach, running a QDA classifier on the extracted descriptors. Results were then aggregated at track level by computing the modal label value among those assigned to the all the ROIs of the same track. Results have been represented in terms of AUC and FROC curve using a LOEO cross-validation strategy. The first row of Table 4 lists the performance achieved also in terms of FPtpBV obtained at sensitivity equal to 0.7, 0.8, and 0.9.

As a second comparative approach, we applied Local Binary Pattern (LBP) image mapping and histogram-based feature extraction [54]. Each ROI was coded through LBP (radius and neighbors were set to the default values 1 and 8, respectively) and the resulting histogram bin heights were used as features. QDA was then used in a LOEO cross-validation procedure to validate the approach. The second row of Table 4 lists the performance achieved.

Results give evidence to the fact that performance achieved by DL architecture and LBP strategy are not as good as those of our proposed approach.

First, the number of regions fed to Alexnet is not sufficiently high for transfer learning to be totally effective.

Second, LBP approach is perfectly suitable in multiresolution applications where information retrieved appear at different scales. In the present scenario, AD pattern does not sufficiently benefit from the multiscale approach since the relevant information mostly lies in the multidirectionality.

In order to enlarge the variety of tests performed to validate our proposed approach, we applied the whole procedure to the contralateral images where no AD loci were reported. Actually, we expect that in case of benign reports for AD, physiological distortions may occur in both breasts and be present even in the contralateral side.

The test led to a number of F_{TP}BV of 2.67, 4.06, and 5.67 for sensitivity values of 0.7, 0.8, and 0.9, respectively.

6.2. Future developments

Future developments will be devoted to the following aspects:

- (i) Enlarge the dataset with more exams: our approach was developed on a dataset of 20 clinical exams. A larger dataset is needed for a better characterization of tracks that could correspond to the path of an AD as well as to allow an improved training performance of the classifier.
- (ii) Include normal exams in the study: in this work, all exams contained ADs identified by experienced radiologist. To effectively test the whole methodology it would be useful to perform feature extraction in truly normal DBT cases of healthy breasts.
- (iii) In future work, expert radiologists will be asked to examine the true positives provided by the CAD especially in presence of multiple-TPs for the same AD. The expert opinion will not only provide a further system assessment but also will guide in algorithm adjusting and optimization.
- (iv) Correlate the results with the histological results: this is maybe the most challenging improvement of this work. Most of the cases where the AD was detected go to biopsy to eliminate the suspect of malignancy. If we could correlate the breast structure to the malignancy of the AD loci, the patient would not be required to do further exams.
- (v) Enlarge the feature set: in the present work, we only considered texture features extracted in the neighborhood of each point along the track. Future investigations will include track features in analogy to kinematics descriptors used in particle tracking analysis.
- (vi) Fully automatize the tracking process of the algorithm: all the methods used here were developed as a prototype. Next steps aim to develop a closed-form algorithm that could be used by radiologists in a typical day-to-day clinical practice.
- (vii) An undoubted improvement will be also related to the possibility to combine information extracted from CC and MLO views. However, problems related to views-correspondence establishment have to be solved accordingly.
- (viii) Finally, further work will be devoted to the overall performance improvement either in terms of increased sensitivity and in a reduction of false positive per breast volume achieved. This is strongly required also in consideration to the fact that, in clinical practice, this number would be added to the false positives generated by microcalcifications and mass detection.

7. Conclusions

The detection of AD loci is a difficult task although AD could represent the earliest manifestation of breast cancer. Given the subtleness and the high malignancy rate of AD, computer algorithms can support the radiologists during the identification task. Since its initial use in clinical practice, in 2011, DBT exam has been used worldwide for screening and for detecting breast cancer. It improved the detection rate of microcalcifications, the classification of masses and it reduced the patients' recall rate. Besides its largely confirmed benefits over digital mammography, no CAD systems for DBT are commercially available yet, nor CAD systems for detecting AD in DBT slices.

In this scenario, this paper presented a methodology to effectively detect AD in DBT exams. The proposed approach explores the information encoded in the tridimensional domain provided by the

DBT slices. After preprocessing, the suspicious spots independently detected in each slice are linked by means of tracks representing the spreading of AD through the breast volume. Regions of interest are extracted in the neighborhood of each tracked point for being analyzed and used as features for classification. The results achieved by this work have proven that our methodology can be used to support radiologists in interpreting DBT slices for the scope of detecting AD.

Ethical approval

Ethics committee approval and informed consent have been collected during DBT acquisition.

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgment

This work was supported by Brazil-Europe ERASMUS MUNDUS (BEMUNDUS) Program, by the Brazilian National Council for Scientific and Technological Development (CNPq) and by the São Paulo Research Foundation (FAPESP), grant #2015/20812-5. The content is solely the responsibility of the authors and does not necessarily represent the official views of the BEMUNDUS, CNPq or FAPESP.

References

- [1] WHO, World Health Organization – Breast Cancer, <http://www.who.int/>, (Accessed: 15.01.17).
- [2] J.G. Elmore, C.Y. Nakano, T.D. Koepsell, L.M. Desnick, C.J. D'Orsi, D.F. Ransohoff, International variation in screening mammography interpretations in community-based programs, *J. Natl. Cancer Inst.* 95 (18) (2003) 1384–1393.
- [3] A.D.A. Maidment, Digital mammography, *Semin. Roentgenol.* 38 (3) (2003) 216–230.
- [4] U. Veronesi, P. Boyle, A. Goldhirsch, R. Orecchia, G. Viale, Breast cancer, *Lancet* 365 (2005) 1727–1741.
- [5] A. Karellas, S. Vedantham, Breast cancer imaging: a perspective for the next decade, *Med. Phys.* 35 (11) (2008) 4878–4897.
- [6] C.G. Glynn, D.M. Farria, B.S. Monsees, J.T. Salcman, K.N. Wiele, C.F. Hildebolt, Effect of transition to digital mammography on clinical outcomes, *Radiology* 260 (3) (2011) 664–670.
- [7] M. Bahl, L.R. Lamb, C.D. Lehman, Pathologic outcomes of architectural distortion on digital 2D versus tomosynthesis mammography, *Am. J. Roentgenol.* 209 (5) (2017) 1162–1167.
- [8] S. Vedantham, A. Karellas, G.R. Vijayaraghavan, D.B. Kopans, Digital breast tomosynthesis: state of art, *Radiology* 277 (3) (2015) 663–684.
- [9] J.A. Baker, J.Y. Lo, Breast tomosynthesis: state-of-the-art and review of the literature, *Acad. Radiol.* 18 (10) (2011) 1298–1310.
- [10] F.J. Gilbert, L. Tucker, K.C. Young, Digital breast tomosynthesis (DBT): a review of the evidence for use as a screening tool, *Clin. Radiol.* 71 (2) (2016) 141–150.
- [11] H.-P. Chan, R.K. Samala, J. Hadjiiski, L.M. an Wei, Computer-aided diagnosis of breast cancer with tomosynthesis, in: P.M. de Azevedo-Marques, A. Mencattini, M. Salmeri, R.M. Rangayyan (Eds.), *Medical Image Analysis and Informatics. Computer-Aided Diagnosis and Therapy*, CRC Press, Taylor & Francis Group, FL, 2018, pp. 241–268 (Chap. 10).
- [12] H.-P. Chan, M.A. Helvie, L. Hadjiiski, D.O. Jeffries, K.A. Klein, C.H. Neal, M. Noroozian, C. Paramagul, M.A. Roubidoux, Characterization of breast masses in digital breast tomosynthesis and digital mammograms: an observer performance study, *Acad. Radiol.* 24 (11) (2017) 1372–1379.
- [13] I. Sechopoulos, A review of breast tomosynthesis. Part I. The image acquisition process, *Med. Phys.* 40 (1) (2013) 014301.
- [14] E.A. Rafferty, J.M. Park, L.E. Philpotts, S.P. Poplack, J.H. Sumkin, E.F. Halpern, L.T. Niklason, Assessing radiologist performance using combined digital mammography and breast tomosynthesis compared with digital mammography alone: results of a multicenter, multireader trial, *Radiology* 266 (1) (2013) 104–113.
- [15] K.M. Ray, E. Turner, E.A. Sickles, B.N. Joe, Suspicious findings at digital breast tomosynthesis occult to conventional digital mammography: imaging features and pathology findings, *Breast J.* 21 (5) (2015) 538–542.
- [16] M.A. Durand, B.M. Haas, X. Yao, J.L. Geisel, M. Raghu, R.J. Hooley, L.J. Horvath, L.E. Philpotts, Early clinical experience with digital breast tomosynthesis for screening mammography, *Radiology* 274 (1) (2015) 85–92.
- [17] T.I. Alshafei, J.V. Nguyen, C.M. Rochman, B.T. Nicholson, J.T. Patrie, J.A. Harvey, Outcome of architectural distortion detected only at breast tomosynthesis versus 2D mammography, *Radiology* 288 (1) (2018) 38–46.

- [18] S. Gaur, V. Dialani, P.J. Slanetz, R.L. Eisenberg, Architectural distortion of the breast, *Am. J. Roentgenol.* 201 (2013) 662–670.
- [19] M. Bahl, J.A. Baker, E.N. Kinsey, S.V. Ghatge, Architectural distortion on mammography: correlation with pathologic outcomes and predictors of malignancy, *Am. J. Roentgenol.* 205 (6) (2015) 1339–1345.
- [20] R.K. Samala, H.-P. Chan, L. Hadjiiski, M.A. Helvie, J. Wei, K. Cha, Mass detection in digital breast tomosynthesis: deep convolutional neural network with transfer learning from mammography, *Med. Phys.* 43 (12) (2016) 6654–6666.
- [21] D.H. Kim, S.T. Kim, J.M. Chang, Y.M. Ro, Latent feature representation with depth directional long-term recurrent learning for breast masses in digital breast tomosynthesis, *Phys. Med. Biol.* 62 (3) (2017) 1009.
- [22] M. Yousefi, A. Krzyżak, C.Y. Suen, Mass detection in digital breast tomosynthesis data using convolutional neural networks and multiple instance learning, *Comput. Biol. Med.* 96 (2018) 283–293.
- [23] C.J. D'Orsi (Ed.), *ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System*, American College of Radiology, 2013.
- [24] J.C. Dean, C.C. Ilvento, Improved cancer detection using computer-aided detection with diagnostic and screening mammography – prospective study of 104 cancers, *Am. J. Roentgenol.* 187 (2006) 20–28.
- [25] S.K. Yang, W.K. Moon, N. Cho, J.S. Park, J.H. Cha, S.M. Kim, S.J. Kim, J.-G. Im, Screening mammography-detected cancers: sensitivity of a computer-aided detection system applied to full-field digital mammograms, *Radiology* 244 (1) (2007) 104–111.
- [26] P. Casti, A. Mencattini, M. Salmeri, A. Ancona, F. Mangeri, M. Pepe, R. Rangayyan, Contour-independent detection and classification of mammographic lesions, *Biomed. Signal Process. Control* 25 (2016) 165–177.
- [27] R.M. Rangayyan, S. Banik, J.E.L. Desautels, Computer-aided detection of architectural distortion in prior mammograms of interval cancer, *J. Digital Imaging* 23 (5) (2010) 611–631.
- [28] W.I. Suleiman, M.F. McEntee, S.J. Lewis, M.A. Rawashdeh, D. Georgian-smith, R. Heard, K. Tapia, P.C. Brennan, In the digital era, architectural distortion remains a challenging radiological task, *Clin. Radiol.* 71 (1) (2016) e35–e40.
- [29] R.A. Benedikt, J.E. Boatsman, C.A. Swann, A.D. Kirkpatrick, A.Y. Toledano, Concurrent computer-aided detection improves reading time of digital breast tomosynthesis and maintains interpretation performance in a multireader multicase study, *Am. J. Roentgenol.* 210 (3) (2018) 685–694.
- [30] L. Morra, D. Sacchetto, M. Durando, S. Agliozzo, L.A. Carbonaro, S. Delsanto, B. Pesce, D. Persano, G. Mariscotti, V. Marra, et al., Breast cancer: computer-aided detection with digital breast tomosynthesis, *Radiology* 277 (1) (2015) 56–63.
- [31] J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J.-Y. Tinevez, D.J. White, V. Hartenstein, K. Eliceiri, P. Tomancak, A. Cardona, Fiji: an open-source platform for biological-image analysis, *Nat. Methods* 9 (2012) 676–682.
- [32] S. Parlato, A. De Ninno, R. Molfetta, E. Toschi, D. Salerno, A. Mencattini, G. Romagnoli, A. Fragale, L. Roccazzello, M. Buoncervello, et al., 3D microfluidic model for evaluating immunotherapy efficacy by tracking dendritic cell behaviour toward tumor cells, *Sci. Rep.* 7 (2017).
- [33] E. Biselli, E. Agliari, A. Barra, F.R. Bertani, A. Gerardino, A. De Ninno, A. Mencattini, D. Di Giuseppe, F. Mattei, G. Schiavoni, et al., Organs on chip approach: a tool to evaluate cancer-immune cells interactions, *Sci. Rep.* 7 (2017) 12737.
- [34] M. Nemoto, S. Honmura, A. Shimizu, D. Furukawa, H. Kobatake, S. Nawano, A pilot study of architectural distortion detection in mammograms based on characteristics of line shadows, *Int. J. Comput. Assist. Radiol. Surg.* 4 (1) (2009) 27–36.
- [35] M. Jasionowska, A. Przelaskowski, A two-step method for detection of architectural distortions in mammograms, *Inf. Technol. Biomed.* 69 (2010) 73–84.
- [36] S. Minavathi, S. Murali, M.S. Dinesh, Model based approach for detection of architectural distortions and spiculated masses in mammograms, *Int. J. Comput. Sci. Eng.* 3 (11) (2011) 3534–3546.
- [37] R.M. Rangayyan, S. Banik, J. Chakraborty, S. Mukhopadhyay, J.E.L. Desautels, Measures of divergence of oriented patterns for the detection of architectural distortion in prior mammograms, *Int. J. Comput. Assist. Radiol. Surg.* 8 (4) (2013) 527–545.
- [38] G. Palma, I. Bloch, S. Muller, Detection of masses and architectural distortions in digital breast tomosynthesis images using fuzzy and a contrario approaches, *Pattern Recogn.* 47 (7) (2014) 2467–2480.
- [39] A. Kamra, V.K. Jain, S. Singh, S. Mittal, Characterization of architectural distortion in mammograms based on texture analysis using support vector machine classifier with clinical evaluation, *J. Digital Imaging* 29 (1) (2016) 104–114.
- [40] X. Liu, L. Zhai, T. Zhu, Z. Yang, Architectural distortion recognition based on a subclass technique and the sparse representation classifier, in: 2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), no. 61403287, IEEE, 2016, pp. 422–426.
- [41] H.C.R. Oliveira, D.R. Moraes, G.A. Reche, L.R. Borges, J.H. Catani, N. de Barros, C.F.E. Melo, A. Gonzaga, M.A.C. Vieira, A new texture descriptor based on local micro-pattern for detection of architectural distortion in mammographic images, in: S.G. Armato, N.A. Petrick (Eds.), *Proc. SPIE Medical Imaging 2017: Computer-Aided Diagnosis*, Vol. 10134, 2017, p. 101342U.
- [42] V.A. Kovalev, M. Petrou, Y.S. Bondar, Texture anisotropy in 3-D images, *IEEE Trans. Image Process.* 8 (3) (1999) 346–360.
- [43] W. Chen, M.L. Giger, H. Li, U. Bick, G.M. Newstead, Volumetric texture analysis of breast lesions on contrast-enhanced magnetic resonance images, *Magn. Reson. Med.* 58 (3) (2007) 562–571.
- [44] R.M. Rangayyan, T.M. Nguyen, F.J. Ayres, A.K. Nandi, Effect of pixel resolution on texture features of breast masses in mammograms, *J. Digital Imaging* 23 (5) (2010) 547–553.
- [45] P. Casti, A. Mencattini, M. Salmeri, A. Ancona, F. Mangeri, M.L. Pepe, R.M. Rangayyan, Estimation of the breast skin-line in mammograms using multidirectional Gabor filters, *Comput. Biol. Med.* 43 (11) (2013) 1870–1881.
- [46] P. Casti, A. Mencattini, M. Salmeri, R.M. Rangayyan, Analysis of structural similarity in mammograms for detection of bilateral asymmetry, *IEEE Trans. Med. Imaging* 34 (2) (2015) 662–671.
- [47] F.J. Ayres, R.M. Rangayyan, Design and performance analysis of oriented feature detectors, *J. Electron. Imaging* 16 (2) (2007), 023007–023007.
- [48] E. Abbena, S. Salamon, A. Gray, *Modern Differential Geometry of Curves and Surfaces with Mathematica*, CRC Press, 2006.
- [49] J.-Y. Tinevez, N. Perry, J. Schindelin, G.M. Hoopes, G.D. Reynolds, E. Laplantine, S.Y. Bednarek, S.L. Shorte, K.W. Eliceiri, Trackmate: an open and extensible platform for single-particle tracking, *Methods* 115 (2017) 80–90.
- [50] K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S.L. Schmid, G. Danuser, Robust single-particle tracking in live-cell time-lapse sequences, *Nat. Methods* 5 (2008) 695–702.
- [51] A. Hamidinekoo, E. Denton, A. Rampun, K. Honnor, R. Zwigelaar, Deep learning in mammography and breast histology, an overview and future trends, *Med. Image Anal.* 47 (2018) 45–67.
- [52] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, *Advances in Neural Information Processing Systems* (2012) 1097–1105.
- [53] S.J. Pan, Q. Yang, et al., A survey on transfer learning, *IEEE Trans. Knowl. Data Eng.* 22 (10) (2010) 1345–1359.
- [54] T. Ojala, M. Pietikainen, T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (7) (2002) 971–987.

7.5 Artigo 5: *SPIE Medical Imaging 2019*

Trabalho apresentado e publicado no congresso internacional
SPIE Medical Imaging (SPIE 2018),
conferência *Computer-Aided Diagnosis*:
<http://doi.org/10.1117/12.2513021>

PROCEEDINGS OF SPIE

[SPIDigitalLibrary.org/conference-proceedings-of-spie](https://spiedigitallibrary.org/conference-proceedings-of-spie)

Exploratory learning with convolutional autoencoder for discrimination of architectural distortion in digital mammography

Helder C. R. de Oliveira, Carlos F. E. Melo, Juliana H. Catani, Nestor de Barros, Marcelo A. da Costa Vieira

Helder C. R. de Oliveira, Carlos F. E. Melo, Juliana H. Catani, Nestor de Barros, Marcelo A. da Costa Vieira, "Exploratory learning with convolutional autoencoder for discrimination of architectural distortion in digital mammography," Proc. SPIE 10950, Medical Imaging 2019: Computer-Aided Diagnosis, 109502F (13 March 2019); doi: 10.1117/12.2513021

SPIE.

Event: SPIE Medical Imaging, 2019, San Diego, California, United States

Exploratory Learning with Convolutional Autoencoder for Discrimination of Architectural Distortion in Digital Mammography

Helder C. R. de Oliveira^a, Carlos F. E. Melo^b, Juliana H. Catani^c,
Nestor de Barros^c, and Marcelo A. da Costa Vieira^{a,*}

^aDept. of Electrical and Computing Engineering, University of São Paulo, São Carlos, Brazil

^bClínica Eco & Mama Diagnóstico Digital, São Carlos, Brazil

^cDept. of Radiology, University of São Paulo, São Paulo, Brazil

heldercro@usp.br, gerencia@ecoemama.com.br

{juliana.catani, nestor.barros}@hc.fm.usp.br, mvieira@sc.usp.br

ABSTRACT

This work presents a deep learning approach based on autoencoder to improve the detection of architectural distortion (AD) in digital mammography. AD can be the earliest sign of breast cancer, appearing before the formation of any mass or calcification. However, it is very difficult to be detected and almost 50% of the cases are missed by the radiologists. Thus, we designed an autoencoder, based on a convolutional neural network (CNN), to work as a feature descriptor in a computer-aided detection (CAD) pipeline with the objective of detecting AD in digital mammography. This model was trained with 140,000 regions-of-interest (ROI) extracted from clinical mammograms. These samples were divided in two groups, with and without AD, according to the radiologist's report. Validation was done comparing the classifier performance when using the proposed autoencoder and other well-known feature descriptors, commonly used for the task of detecting AD in digital mammograms. The results showed that the performance of the autoencoder is slightly higher than that of other descriptors. However, the complexity and the computational cost of the autoencoder is much higher when compared to the hand-crafted descriptors.

Keywords: architectural distortion, breast cancer, digital mammography, feature learning, autoencoder, deep learning.

1. INTRODUCTION

The detection of breast lesions in mammography is fundamental to the correct diagnosis and necessary to the following up treatment of breast cancer.¹ Breast cancer is the main cause of death by cancer among women worldwide.¹ The treatment is up to 30% more effective when the disease is diagnosed in earliest stage.²⁻⁴

The analysis of a mammogram is time consuming and expensive. In order to assist radiologist in this task, Computer-Aided Detection (CAD) softwares have been used in clinical practice to indicate suspicious regions where the radiologist should pay more attention, in such way providing a “second opinion” to the physician.^{4,5} CAD systems often use computer vision and pattern recognition techniques to analyze the image and find patterns which look suspicious.⁵⁻⁷

When reading mammograms, the radiologist try to identify potential sites of calcifications, masses and/or architectural distortions.^{6,8} Among these, Architectural Distortion (AD) is considered the earliest manifestation of a possible breast cancer.^{9,10} In the case of false-negative mammograms, AD is the most common finding and about 50% of the cases are missed.¹¹ This happens because it subtly changes the breast parenchyma in such way that it does not have a mass associated nor visible border.^{9,12}

Several studies have been proposed to detect AD in mammography images.^{7,11,13-16} Most of them use hand-crafted texture descriptors, developed for general proposes, in a Computer-aided Detection (CAD) pipeline.⁴

*Corresponding author: M.A.C.V.: mvieira@sc.usp.br

Autoencoder is a model for *unsupervised learning* where a bank of filters are trained to efficiently represent the input data by means of a code. When properly trained, the autoencoder can be used as a feature descriptor and can be connected to any classifier.^{17,18} In this study, we designed an autoencoder, based on a convolutional neural network (CNN), to work as a feature descriptor in a CAD pipeline with the objective of detecting AD in digital mammography.

2. MATERIALS AND METHODS

An overview of the pipeline used in this work to train and validate the proposed autoencoder is presented in Figure 1. Each part is explained in the next sections.

Figure 1. Overview of the proposed pipeline.

2.1 Digital Mammography Datasets

Two datasets without any overlapping data of clinical Full-field Digital Mammography (FFDM) images were created, Dataset #1 and Dataset #2. The Dataset #1 was used to train the models, while the second one was used to the validation step. In both sets the AD cases were confirmed by clinical report and manually segmented by a radiologist with more than 14 years of experience reading mammograms.

The Dataset #1 is composed of 350 cases, 175 with AD and 175 without any breast lesion. Images were collected at the Institute of Radiology (InRAD) of the University of São Paulo (USP), under Institutional Review Board (IRB) approval*. Figure 2, shows two images where the AD location, marked by the radiologist, is highlighted by a red circle.

Since AD is well characterized by its neighborhood, given a central point in the segmented area, 50 regions-of-interest (ROI) of 256×256 pixels were extracted from each image. In the case of mammograms with AD, the selection of ROIs was performed considering sliding windows that contained the center of the region marked by the radiologist, as shown in Figure 2b. The yellow star represents the center of the AD region and the dashed rectangles represent the selected ROIs, considered as positive for AD. For the normal cases, other set of 50 ROIs were taken randomly from each mammogram considered as normal. After the ROIs selection, the Dataset #1 was composed of 17,500 samples: 175 images $\times$ 2 classes (AD and not AD) $\times$ 50 ROIs/image.

The Dataset #2 was composed by 52 cases, 26 with AD and 26 without, acquired in another imaging center, Eco & Mama clinic, under IRB approval[†]. It was used for the validation step of this work.

*CAAE #56699016.7.0000.0065

[†]CAAE #80264717.3.0000.5393

Figure 2. Examples of clinical mammograms used in this study, (a) MLO and (b) CC views. The red circle in (a) represents the AD marked by the radiologist. The dashed squares in the image (b) represent ROIs extracted from the AD region.

2.2 Data Augmentation

As mentioned before, the Dataset #1 was used for training. Autoencoder is based on a deep convolutional network and it is necessary a large amount of data for the training step to avoid over-fitting. Thus, we use the data augmentation approach to increase our database.^{18,19} We applied seven types of geometric transformations (combining image rotation and flip) to existing data, as shown in Figure 3, to create a database of approximately 140,000 samples.

Figure 3. Example of geometric transformations which were applied to existing data to increase our image database (data augmentation): (a) original ROI. (b-d) after applying rotation of 90° , 180° and 270° . (e-h) Horizontal flip applied to ROIs showed in (a-d), respectively.

2.3 Convolutional Autoencoders

Representation learning is a machine learning approach where the aim is to build artificial neural networks capable of learn features from the training data. *Convolutional autoencoders* are the emergent set of algorithms.²⁰ The autoencoder model, exemplified in Figure 4, is usually composed of two parts, the encoder $f(\cdot)$ and the decoder $g(\cdot)$. The encoder has the objective of decomposing the input image through convolutional layers, in order to extract relevant features and reduce dimensionality.¹⁹ The decoder will try to reconstruct the input image based only on the code generated by the encoder.

Figure 4. Autoencoder setup, which is composed of the encoder (darker blue) and the decoder (lighter blue), along with its layers (convolutional, normalization, max. pooling and up sampling). The numbers inside each box (layer) shows the dimension of the output data. The yellow box represents the “code” generated by the model to describe the input image.

The training step of the autoencoder design is unsupervised, i.e., we do not use labeled data.¹⁸ Thereby, the training algorithm basically try to minimize the Mean Squared Error (MSE) between the input and output samples, as shown in Equation 1:

$$\min \left[\frac{1}{n} \sum_{i=1}^n |g(f(x_i)) - x_i|^2 \right] \quad (1)$$

where, f is the encoder, g is the decoder, n is the amount of samples and x_i is the i -esimal sample.

Once that autoencoders can learn how to represent the input data by means of a code, it can be trained to be used as a specialized feature descriptor.¹⁷ The Figure 4, shows on the left some examples of input images used in this work, as well as its respective reconstructed version after training the model (on the right).

Each convolutional layer of the proposed autoencoder has the kernel size of 3×3 , hyperbolic tangent as activation function and the padding method keeps the output data size equal to the input. Other parameters: batch size = 10, training epochs = 500, train/validation size = 80%/20% and optimizer = **rmsprop**.

2.4 Multilayer Perceptron Network

After the training step, we can remove the encoder part to plug it to a classifier, as shown in the Figure 4. In this work, we used a Multilayer Perceptron (MLP) model with three neural layers composed of 100, 50 and 1 neurons, respectively, as a classifier. For the first two layers the activation function is the *Rectified Linear Unit* (ReLU) followed by the last layer with a sigmoid function.

Since the last layer of the encoder is a volume ($16 \times 16 \times 16$), the output data needs to be flattened in a row vector of 4,096 values. This is performed by adding an auxiliary layer for this purpose, as shown in the green box of Figure 5. When complete, the model (encoder + flatten layer + MLP) needs to be trained. However, at this time, the optimization will occur only in the neural layers. To avoid new changes in the already trained encoder, its layers need to be “frozen”. This task was performed, for 1,000 epochs, using the Dataset #1.

Figure 5. Encoder, taken from the model presented at Figure 4, attached to three fully connected neural layers for classification.

3. RESULTS

The performance of the proposed classification pipeline using the autoencoder was compared to other well-known image descriptors: Haralick's Energy and Local Binary Pattern (LBP). Both had good results in the task of AD description as reported in a previous work.¹⁵ Dataset #1 was used to train the Encoder+MLP model, and it was also used to train two new MLP models, one for each descriptor, with the same setup as described before.

The validation of the descriptors was performed using the Dataset #2. The results, in terms of Accuracy and Area Under the Receiver Operating Characteristics Curve (AUC ROC), are summarized in Table 1.

Table 1. Results regarding the evaluation of each descriptor over two different datasets. The best result of each column is highlighted in bold.

	Dataset #1		Dataset #2	
Descriptor	Accuracy	AUC	Accuracy	AUC
Encoder	0.7052	0.7706	0.7535	0.8295
Energy	0.6908	0.7387	0.7473	0.8024
LBP	0.6471	0.7007	0.5000	0.5814

Based on Table 1, considering only the columns related to the Dataset #1, it is possible to see that the accuracy and AUC achieved by the Energy descriptor and the encoder are very close. However, we must be aware that the results of Encoder are related to the samples which were already learned by the model when the autoencoder was trained. In this case, we believe the results regarding Energy and LBP are more expressive, since these methods require only one training step.

When used the Dataset #2, all methods, except the LBP, performed better than for the Dataset #1. ROC analysis of the descriptors and classifiers performance can be seen in Figure 6. For the Dataset #1 (Figure 6a) the Encoder and Haralick's Energy descriptor achieved similar performance. The LBP descriptor achieved the worst performance among them.

Considering Dataset #2 (Figure 6b), the performance of the Encoder and Haralick's Energy descriptors have increased, in comparison to the Dataset #1. However, for the LBP descriptor, the performance has decreased. Encoder and Haralick's Energy descriptors presented similar performance. Encoder achieved better sensitivity and Energy better specificity.

Figure 6. ROC curves related to (a) training using Dataset #1 and, (b) the validation step using Dataset #2.

4. CONCLUSIONS

This paper evaluated a new approach to detect architectural distortion of breast in digital mammography images using the autoencoder technique. The performance of the autoencoder was compared to two other descriptors (LBP and the Haralick's Energy) which presented better results in the task of AD detection, as reported in previous work.¹⁵ The evaluation was performed with two datasets of clinical images without any overlapping sample and collected in two different imaging centers.

The autoencoder was able to learn how to describe specific features in digital images. In this work, it was optimized to the specific task of describing AD in digital mammography. This is different of the concept behind most of hand-crafted descriptors, which were developed to describe general objects in natural images, which might not be as subtle as the AD.

The results shown that proposed autoencoder slightly over performed the other descriptors evaluated with: accuracy of 0.7052, AUC of 0.7706 for the Dataset #1 and, accuracy of 0.7535 and AUC of 0.8295 for Dataset #2. However these results were very close to those ones presented by Haralick's Energy: accuracy of 0.6908, AUC of 0.7387 and, accuracy of 0.7476, AUC of 0.8024, for datasets #1 and #2, respectively. Although the results showed that the performance of the autoencoder is slightly higher when compared to the other image descriptors, the computational cost is much higher for the autoencoder case. Considering our simulations, it took two weeks to finish the autoencoder design (training the autoencoder and the encoder with the MLP) while applying the Haralick's Energy descriptor and training the MLP took about eight hours.

ACKNOWLEDGMENTS

This project was supported by the National Council for Scientific and Technological Development (CNPq).

REFERENCES

- [1] WHO, "World health organization - breast cancer." <http://www.who.int/>. (Accessed: 10 July 2018).
- [2] Elmore, J. G., Nakano, C. Y., Koepsell, T. D., Desnick, L. M., D'Orsi, C. J., and Ransohoff, D. F., "International variation in screening mammography interpretations in community-based programs," *J. Natl. Cancer Inst.* **95**(18), 1384–1393 (2003).

- [3] Veronesi, U., Boyle, P., Goldhirsch, A., Orecchia, R., and Viale, G., "Breast Cancer," *The Lancet* **365**, 1727–1741 (2005).
- [4] Russo, P., ed., [*Handbook of X-ray Imaging: Physics and Technology*], Series in Medical Physics and Biomedical Engineering, CRC Press, Boca Raton, 1 ed. (2017).
- [5] Doi, K., "Computer-aided diagnosis in medical imaging: Historical review, current status and future potential," *Computerized Medical Imaging and Graphics* **31**(4), 198–211 (2007).
- [6] Giger, M. L., Chan, H.-P., and Boone, J., "Anniversary paper: History and status of cad and quantitative image analysis: The role of medical physics and aapm," *Medical Physics* **35**(12), 5799–5820 (2008).
- [7] Oliveira, H. C. R., Mencattini, A., Casti, P., Catani, J. H., de Barros, N., Gonzaga, A., Martinelli, E., and Vieira, M. A. C., "A cross-cutting approach for tracking architectural distortion locii on digital breast tomosynthesis slices," *Biomedical Signal Processing and Control* **50**, 92 – 102 (2019).
- [8] Baker, J. A., Rosen, E. L., Lo, J. Y., Gimenez, E. I., Walsh, R., and Soo, M. S., "Computer-aided detection (CAD) in screening mammography: Sensitivity of commercial CAD systems for detecting architectural distortion," *American Journal of Roentgenology* **181**(4), 1083–1088 (2003).
- [9] Gaur, S., Dialani, V., Slanetz, P. J., and Eisenberg, R. L., "Architectural Distortion of the Breast," *American Journal of Roentgenology* **201**, 662–670 (2013).
- [10] Bahl, M., Baker, J. A., Kinsey, E. N., and Ghate, S. V., "Architectural Distortion on Mammography: Correlation With Pathologic Outcomes and Predictors of Malignancy," *American Journal of Roentgenology* **205**(6), 1339–1345 (2015).
- [11] Suleiman, W. I., McEntee, M. F., Lewis, S. J., Rawashdeh, M. A., Georgian-smith, D., Heard, R., Tapia, K., and Brennan, P. C., "In the digital era, architectural distortion remains a challenging radiological task," *Clinical Radiology* **71**(1), e35–e40 (2016).
- [12] D'Orsi, C. J., ed., [*ACR BI-RADS® Atlas, Breast Imaging Reporting and Data System*], American College of Radiology (2013).
- [13] Rangayyan, R. M., Banik, S., and Desautels, J. E. L., "Computer-Aided Detection of Architectural Distortion in Prior Mammograms of Interval Cancer," *Journal of Digital Imaging* **23**(5), 611–631 (2010).
- [14] Banik, S., Rangayyan, R. M., and Desautels, J. E. L., "Measures of angular spread and entropy for the detection of architectural distortion in prior mammograms," *International Journal of Computer Assisted Radiology and Surgery* **8**(1), 121–134 (2013).
- [15] Oliveira, H. C. R., Moraes, D. R., Reche, G. A., Borges, L. R., Catani, J. H., de Barros, N., Melo, C. F. E., Gonzaga, A., and Vieira, M. A. C., "A new texture descriptor based on local micro-pattern for detection of architectural distortion in mammographic images," in [*Proc.SPIE*], **10134**, 10134–10134–9 (2017).
- [16] Oliveira, H. C. R., Mencattini, A., Casti, P., Martinelli, E., di Natale, C., Catani, J. H., de Barros, N., Melo, C. F. E., Gonzaga, A., and Vieira, M. A. C., "Reduction of false-positives in a cad scheme for automated detection of architectural distortion in digital mammography," in [*Proc.SPIE*], **10575**, 10575–10575–10 (2018).
- [17] Wang, S., Ding, Z., and Fu, Y., "Feature Selection Guided Auto-Encoder," in [*Proc. of the 31th Conference on Artificial Intelligence (AAAI 2017)*], 2725–2731 (2017).
- [18] James, G., Witten, D., Hastie, T., and Tibshirani, R., [*An Introduction to Statistical Learning: with Applications in R*], Springer Texts in Statistics, Springer New York (2013).
- [19] Géron, A., [*Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*], O'Reilly Media (2017).
- [20] Chartrand, G., Cheng, P. M., Vorontsov, E., Drozdal, M., Turcotte, S., Pal, C. J., Kadoury, S., and Tang, A., "Deep learning: A primer for radiologists," *RadioGraphics* **37**(7), 2113–2131 (2017).

Referências

- ABBENA, E.; SALAMON, S.; GRAY, A. **Modern differential geometry of curves and surfaces with Mathematica**. [S.l.]: CRC press, 2006.
- AFSHAR, P.; MOHAMMADI, A.; PLATANIOTIS, K. N.; OIKONOMOU, A.; BENALI, H. From Handcrafted to Deep-Learning-Based Cancer Radiomics: Challenges and opportunities. **IEEE Signal Processing Magazine**, 2019. v. 36, n. 4, p. 132–160, 2019.
- ANAND, S.; RATHANA, R. A. V. Architectural Distortion Detection in Mammogram using Contourlet Transform and Texture Features. **International Journal of Computer Applications**, 2013. v. 74, n. 5, p. 12–19, 2013.
- AYRES, F.; RANGAYYAN, R. M. Characterization of architectural distortion in mammograms. In: **Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439)**. [S.l.]: IEEE, 2003. p. 886–889.
- AYRES, F. J.; RANGAYYAN, R. M. Design and performance analysis of oriented feature detectors. **Journal of Electronic Imaging**, 2007. v. 16, n. 2, 2007.
- BAHL, M.; BAKER, J. A.; KINSEY, E. N.; GHATE, S. V. Architectural Distortion on Mammography: Correlation With Pathologic Outcomes and Predictors of Malignancy. **American Journal of Roentgenology**, 2015. American Roentgen Ray Society, v. 205, n. 6, p. 1339–1345, 2015.
- BAKER, J. A.; LO, J. Y. Breast Tomosynthesis: State-of-the-Art and Review of the Literature. **Academic Radiology**, 2011. v. 18, n. 10, p. 1298–1310, 2011.
- BAKER, J. A.; ROSEN, E. L.; LO, J. Y.; GIMENEZ, E. I.; WALSH, R.; SOO, M. S. Computer-aided detection (CAD) in screening mammography: Sensitivity of commercial CAD systems for detecting architectural distortion. **American Journal of Roentgenology**, 2003. v. 181, n. 4, p. 1083–1088, 2003.
- BANIK, S.; RANGAYYAN, R. M.; DESAUTELS, J. E. L. Detection of Architectural Distortion in Prior Mammograms. **IEEE Transactions on Medical Imaging**, 2011. v. 30, n. 2, p. 279–294, 2011.
- BANIK, S.; RANGAYYAN, R. M.; DESAUTELS, J. E. L. Measures of angular spread and entropy for the detection of architectural distortion in prior mammograms.

International Journal of Computer Assisted Radiology and Surgery, 2013. v. 8, n. 1, p. 121–134, 2013.

BISWAS, S. K.; MUKHERJEE, D. P. Recognizing Architectural Distortion in Mammogram: A Multiscale Texture Modeling Approach with GMM. **IEEE Transactions on Biomedical Engineering**, 2011. v. 58, n. 7, p. 2023–2030, 2011.

BOAVENTURA, I.; GONZAGA, A. Edge detection in digital images using fuzzy numbers. **International Journal of Innovative Computing and Applications**, 2009. v. 2, n. 1, p. 1–12, 2009.

BOZEK, J.; DELAC, K.; GRGIC, M. Computer-aided detection and diagnosis of breast abnormalities in digital mammography. **ELMAR, 2008. 50th International Symposium**, 2008. v. 1, p. 45–52, 2008.

CASTI, P.; MENCATTINI, A.; SALMERI, M.; ANCONA, A.; MANGERI, F.; PEPE, M. L.; RANGAYYAN, R. M. Estimation of the breast skin-line in mammograms using multidirectional Gabor filters. **Computers in Biology and Medicine**, 2013. v. 43, n. 11, p. 1870–1881, 2013.

CASTI, P.; MENCATTINI, A.; SALMERI, M.; ANCONA, A.; MANGERI, F.; PEPE, M.; RANGAYYAN, R. Contour-independent detection and classification of mammographic lesions. **Biomedical Signal Processing and Control**, 2016. v. 25, p. 165–177, 2016.

CHIERICI, C. E. O.; VIEIRA, R. T.; FERRAZ, C. T.; TRAVAINI, J.; GONZAGA, A. A New Approach for Analyzing Rotated Textures. In: **Anais do IX Workshop de Visão Computacional**. [S.l.: s.n.], 2013.

DAVIES, E. **Computer and Machine Vision: Theory, Algorithms, Practicalities**. [S.l.]: Elsevier, 2012.

DEAN, J. C.; ILVENTO, C. C. Improved Cancer Detection Using Computer-Aided Detection with Diagnostic and Screening Mammography - prospective study of 104 cancers. **American Journal of Roentgenology**, 2006. v. 187, p. 20–28, 2006.

DIBBLE, E. H.; LOURENCO, A. P.; BAIRD, G. L.; WARD, R. C.; MAYNARD, A. S.; MAINIERO, M. B. Comparison of digital mammography and digital breast tomosynthesis in the detection of architectural distortion. **European Radiology**, 2018. v. 28, n. 1, p. 3–10, 2018.

DOI, K. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. **Computerized Medical Imaging and Graphics**, 2007. v. 31, n. 4, p. 198 – 211, 2007.

DURAND, M. A.; HAAS, B. M.; YAO, X.; GEISEL, J. L.; RAGHU, M.; HOOLEY, R. J.; HORVATH, L. J.; PHILPOTTS, L. E. Early Clinical Experience with Digital Breast Tomosynthesis for Screening Mammography 1. **Radiology**, 2015. v. 274, n. 1, 2015.

ELMORE, J. G.; NAKANO, C. Y.; KOEPESELL, T. D.; DESNICK, L. M.; D'ORSI, C. J.; RANSOHOFF, D. F. International variation in screening mammography interpretations in community-based programs. **J. Natl. Cancer Inst.**, 2003. v. 95, n. 18, p. 1384–1393, 2003.

FERNÁNDEZ-DELGADO, M.; CERNADAS, E.; BARRO, S.; AMORIM, D. Do we need hundreds of classifiers to solve real world classification problems? **Journal of Machine Learning Research**, 2014. v. 15, p. 3133–3181, 2014.

FERRARI, R.; RANGAYYAN, R.; DESAUTELS, J.; BORGES, R.; FRERE, A. Automatic Identification of the Pectoral Muscle in Mammograms. **IEEE Transactions on Medical Imaging**, 2004. v. 23, n. 2, p. 232–245, 2004.

FERRAZ, C. T.; PEREIRA, O.; GONZAGA, A. Feature description based on center-symmetric local mapped patterns. In: **Proceedings of the 29th Annual ACM Symposium on Applied Computing - SAC '14**. New York, New York, USA: ACM Press, 2014. (SAC '14), p. 39–44.

FERRAZ, C. T.; PEREIRA, O.; ROSA, M. V.; GONZAGA, A. Object Recognition Based on Bag of Features and a New Local Pattern Descriptor. **International Journal of Pattern Recognition and Artificial Intelligence**, 2014. World Scientific Publishing Co., v. 28, n. 08, p. 1455010, 2014.

GAUR, S.; DIALANI, V.; SLANETZ, P. J.; EISENBERG, R. L. Architectural Distortion of the Breast. **American Journal of Roentgenology**, 2013. v. 201, p. 662–670, 2013.

GLYNN, C. G.; FARRIA, D. M.; MONSEES, B. S.; SALCMAN, J. T.; WIELE, K. N.; HILDEBOLT, C. F. Effect of transition to digital mammography on clinical outcomes. **Radiology**, 2011. v. 260, n. 3, p. 664–670, 2011.

GONZALEZ, R. C.; WOODS, R. E. **Digital image processing**. [S.l.]: Prentice Hall, 2008.

GORELIK, R.; GAUTREAU, A. Quantitative and unbiased analysis of directional persistence in cell migration. **Nature protocols**, 2014. Nature Publishing Group, v. 9, n. 8, p. 1931–43, 2014.

GREENSPAN, H.; GINNEKEN, B. van; SUMMERS, R. M. Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique. **IEEE Transactions on Medical Imaging**, 2016. v. 35, n. 5, p. 1153–1159, 2016.

GUO, X.; LIU, X.; ZHU, E.; YIN, J. Deep clustering with convolutional autoencoders. In: LIU, D.; XIE, S.; LI, Y.; ZHAO, D.; EL-ALFY, E.-S. M. (Ed.). **Neural Information Processing**. [S.l.]: Springer International Publishing, 2017. p. 373–382.

HANDA, T.; ZHANG, X.; HOMMA, N.; ISHIBASHI, T.; KAWASUMI, Y.; ABE, M.; SUGITA, N.; YOSHIZAWA, M. DoG-Based Detection of Architectural Distortion in Mammographic Images for Computer-Aided Detection. 2012. p. 762–767, 2012.

HARALICK, R. M. Statistical and structural approaches to texture. In: **Proceedings of the IEEE**. [S.l.: s.n.], 1979. v. 67, n. 5, p. 786–804.

HEATH, M.; BOWYER, K.; KOPANS, D.; MOORE, R.; KEGELMEYER, W. P. The digital database for screening mammography. In: **Proceedings of the 5th international workshop on digital mammography**. [S.l.: s.n.], 2000. p. 212–218.

- HEIKKILÄ, M.; PIETIKÄINEN, M.; SCHMID, C. Description of Interest Regions with Center-Symmetric Local Binary Patterns. In: KALRA, P. K.; PELEG, S. (Ed.). **Computer Vision, Graphics and Image Processing**. [S.l.]: Springer Berlin Heidelberg, 2006, (Lecture Notes in Computer Science, v. 4338). p. 58–69.
- HOLOGIC. **Understanding ImageChecker® CAD 10.0 - User Guide**. [S.l.], 2014.
- INCA. Instituto nacional de câncer josé alencar gomes da silva. 2019. 2019. Disponível em: <<http://www.inca.gov.br/>>. Acesso em: 5 de abr. de 2019.
- JAIN, A. K.; FARROKHNI, F. Unsupervised texture segmentation using Gabor filters. **Pattern Recognition**, 1991. v. 24, n. 12, p. 1167–1186, jan 1991.
- JASIONOWSKA, M.; PRZELASKOWSKI, A. A two-step method for detection of architectural distortions in mammograms. **Information Technology in Biomedicine**, 2010. v. 69, p. 73–84, 2010.
- KAMRA, A.; JAIN, V. K.; SINGH, S.; MITTAL, S. Characterization of Architectural Distortion in Mammograms Based on Texture Analysis Using Support Vector Machine Classifier with Clinical Evaluation. **Journal of Digital Imaging**, 2016. v. 29, n. 1, p. 104–114, 2016.
- KARELLAS, A.; VEDANTHAM, S. Breast cancer imaging: a perspective for the next decade. **Medical Physics**, 2008. v. 35, n. 11, p. 4878–4897, Nov 2008.
- KOHLI, A.; JHA, S. Why CAD Failed in Mammography. **Journal of the American College of Radiology**, 2018. American College of Radiology, v. 15, n. 3, p. 535–537, 2018.
- LAKHANI, P.; GRAY, D. L.; PETT, C. R.; NAGY, P.; SHIH, G. Hello World Deep Learning in Medical Imaging. **Journal of Digital Imaging**, 2018. Journal of Digital Imaging, v. 31, n. 3, p. 283–289, 2018.
- LAKSHMANAN, R.; P, S. T.; THOMAS, V.; JACOB, S. M.; PRATAB, T. Detection of Architectural Distortion in Mammogram. **Journal of Image and Graphics**, 2015. v. 2, n. 2, p. 123–127, 2015.
- LAKSHMANAN, R.; T.P., S.; THOMAS, V.; JACOB, S. M.; PRATAB, T. A Preprocessing Method for Reducing Search Area for Architectural Distortion in Mammographic Images. In: **2014 Fourth International Conference on Advances in Computing and Communications**. [S.l.: s.n.], 2014. p. 101–104.
- LITJENS, G.; KOOI, T.; BEJNORDI, B. E.; SETIO, A. A. A.; CIOMPI, F.; GHAFORIAN, M.; LAAK, J. A. van der; GINNEKEN, B. van; SÁNCHEZ, C. I. A survey on deep learning in medical image analysis. **Medical Image Analysis**, 2017. v. 42, p. 60–88, 2017.
- LIU, X.; ZHAI, L.; ZHU, T.; ZHANG, K. A new feature selection method for the detection of architectural distortion in mammographic images. In: FALCO, C. M.; JIANG, X. (Ed.). **Eighth International Conference on Digital Image Processing (ICDIP 2016)**. [S.l.: s.n.], 2016. v. 10033.

- LIU, X.; ZHAI, L.; ZHU, T.; YANG, Z. Architectural distortion recognition based on a subclass technique and the sparse representation classifier. In: **2016 9th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)**. [S.l.]: IEEE, 2016. p. 422–426.
- LIU, X.; ZHAI, L.; ZHU, T.; LIU, J.; ZHANG, K.; HU, W. Multiple TBSVM-RFE for the detection of architectural distortion in mammographic images. **Multimedia Tools and Applications**, 2017. Multimedia Tools and Applications, p. 1–30, 2017.
- MAIDMENT, A. D. A. Digital mammography. **Seminars in Roentgenology**, 2003. v. 38, n. 3, p. 216–230, 2003.
- MAYO, R. C.; LEUNG, J. Artificial intelligence and deep learning – Radiology’s next frontier? **Clinical Imaging**, 2018. Elsevier, v. 49, n. July 2017, p. 87–88, 2018.
- MEIJERING, E.; SMAL, I.; DANUSER, G. Tracking in molecular bioimaging. **IEEE Signal Processing Magazine**, 2006. v. 23, n. 3, p. 46–53, 2006.
- MENCATTINI, A.; SALMERI, M.; RABOTTINO, G.; MEMBER, S.; SALICONE, S.; MEMBER, S. Metrological Characterization of a CADx System for the Classification of Breast Masses in Mammograms. **IEEE Transactions on Instrumentation and Measurement**, 2010. v. 59, n. 11, p. 2792–2799, 2010.
- MINA, L. M.; ISA, N. A. M. A Review of Computer-Aided Detection and Diagnosis of Breast Cancer in Digital Mammography. **Journal of Medical Sciences**, 2015. v. 15, n. 3, p. 110–121, 2015.
- MINAVATHI, S.; MURALI, S.; DINESH, M. S. Model based approach for Detection of Architectural Distortions and Spiculated Masses in Mammograms. **International Journal on Computer Science and Engineering**, 2011. v. 3, n. 11, p. 3534–3546, 2011.
- MOHAMMADI, E.; FATEMIZADEH, E.; SHEIKHZADEH, H.; KHOUBANI, S. A textural approach for recognizing architectural distortion in mammograms. In: **2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP)**. [S.l.: s.n.], 2013. p. 136–140.
- MORRA, L.; COCCIA, N.; CERQUITELLI, T. Optimization of computer aided detection systems: An evolutionary approach. **Expert Systems with Applications**, 2018. v. 100, p. 145–156, 2018.
- MORRA, L.; DELSANTO, S.; GRAZZINI, G. Computer-Aided Diagnosis for X-ray Imaging. In: RUSSO, P. (Ed.). **Handbook of X-ray Imaging: Physics and Technology**. 1. ed. [S.l.]: CRC Press, 2018. cap. 59, p. 1211–1230.
- MUSTRA, M.; GRGIC, M.; RANGAYYAN, R. M. Review of recent advances in segmentation of the breast boundary and the pectoral muscle in mammograms. **Medical & Biological Engineering & Computing**, 2016. v. 54, n. 7, p. 1003–1024, 2016.
- NEMOTO, M.; HONMURA, S.; SHIMIZU, A.; FURUKAWA, D.; KOBATAKE, H.; NAWANO, S. A pilot study of architectural distortion detection in mammograms based on characteristics of line shadows. **International Journal of Computer Assisted Radiology and Surgery**, 2009. v. 4, n. 1, p. 27–36, 2009.

NETPRASAT, O.; AUEPHANWIRIYAKUL, S.; THEERA-UMPON, N. Architectural Distortion Detection from Mammograms Using Support Vector Machine. In: **International Joint Conference on Neural Networks**. [S.l.: s.n.], 2014. p. 3258–3264.

OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on feature distributions. **Pattern Recognition**, 1996. v. 29, n. 1, p. 51–59, 1996.

OJALA, T.; PIETIKAINEN, M.; MAENPAA, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, 2002. v. 24, n. 7, p. 971–987, 2002.

OLIVEIRA, H. C.; MELO, C. F. E.; CATANI, J. H.; BARROS, N.; VIEIRA, M. A. C. Exploratory learning with convolutional autoencoder for discrimination of architectural distortion in digital mammography. In: HAHN, H. K.; MORI, K. (Ed.). **Medical Imaging 2019: Computer-Aided Diagnosis**. [S.l.]: SPIE, 2019. p. 8.

OLIVEIRA, H. C. R.; MENCATTINI, A.; CASTI, P.; MARTINELLI, E.; Di Natale, C.; CATANI, J. H.; BARROS, N.; MELO, C. F.; GONZAGA, A.; VIEIRA, M. A. C. Reduction of false-positives in a CAD scheme for automated detection of architectural distortion in digital mammography. In: MORI, K.; PETRICK, N. (Ed.). **Medical Imaging 2018: Computer-Aided Diagnosis**. [S.l.]: SPIE, 2018. v. 10575, p. 96.

OLIVEIRA, H. C. R.; MENCATTINI, A.; CASTI, P.; CATANI, J. H.; BARROS, N.; GONZAGA, A.; MARTINELLI, E.; Costa Vieira, M. A. A cross-cutting approach for tracking architectural distortion locii on digital breast tomosynthesis slices. **Biomedical Signal Processing and Control**, 2019. v. 50, p. 92–102, 2019.

OLIVEIRA, H. C. R.; MORAES, D. R.; GONZAGA, A.; VIEIRA, M. A. C. Detection of architectural distortion in digitized screen-film mammograms using texture descriptors. In: **Proc. of XII Workshop de Visao Computacional (WVC 2016)**. Campo Grande - MS: [s.n.], 2016. p. 260–265.

OLIVEIRA, H. C. R.; MORAES, D. R.; RECHE, G. A.; BORGES, L. R.; CATANI, J. H.; BARROS, N. de; MELO, C. F. E.; GONZAGA, A.; VIEIRA, M. A. C. A new texture descriptor based on local micro-pattern for detection of architectural distortion in mammographic images. In: ARMATO, S. G.; PETRICK, N. A. (Ed.). **Proc. SPIE Medical Imaging 2017: Computer-Aided Diagnosis**. [S.l.: s.n.], 2017. v. 10134, p. 101342U.

PALMA, G.; BLOCH, I.; MULLER, S. Detection of masses and architectural distortions in digital breast tomosynthesis images using fuzzy and a contrario approaches. **Pattern Recognition**, 2014. v. 47, n. 7, p. 2467–2480, 2014.

PARTYKA, L.; LOURENCO, A. P.; MAINIERO, M. B. Detection of mammographically occult architectural distortion on digital breast tomosynthesis screening: Initial clinical experience. **American Journal of Roentgenology**, 2014. v. 203, n. 1, p. 216–222, 2014.

PHILPOTTS, L.; HOOLEY, R. **Breast Tomosynthesis**. [S.l.]: Elsevier Health Sciences, 2018.

- PIETIKÄINEN, M.; HADID, A.; ZHAO, G.; AHONEN, T. **Computer Vision Using Local Binary Patterns**. [S.l.]: Springer London, 2011. (Computational Imaging and Vision).
- PUJARA, A. C.; HUI, J.; WANG, L. C. Architectural distortion in the era of digital breast tomosynthesis: outcomes and implications for management. **Clinical Imaging**, 2019. Elsevier, v. 54, p. 133–137, 2019.
- RAFFERTY, E. A.; PARK, J. M.; PHILPOTTS, L. E.; POPLACK, S. P.; SUMKIN, J. H.; HALPERN, E. F.; NIKLASON, L. T. Assessing Radiologist Performance Using Combined Digital Mammography and Breast Tomosynthesis Compared with Digital Mammography Alone: Results of a Multicenter, Multireader Trial. **Radiology**, 2013. v. 266, n. 1, p. 104–113, 2013.
- RANGAYYAN, R. M.; AYRES, F. J. Gabor filters and phase portraits for the detection of architectural distortion in mammograms. **Med Bio Eng Comput**, 2006. v. 44, n. 123, p. 883–894, 2006.
- RANGAYYAN, R. M.; BANIK, S.; CHAKRABORTY, J.; MUKHOPADHYAY, S.; DESAUTELS, J. E. L. Measures of divergence of oriented patterns for the detection of architectural distortion in prior mammograms. **International Journal of Computer Assisted Radiology and Surgery**, 2013. v. 8, n. 4, p. 527–545, 2013.
- RANGAYYAN, R. M.; BANIK, S.; DESAUTELS, J. E. L. Computer-Aided Detection of Architectural Distortion in Prior Mammograms of Interval Cancer. **Journal of Digital Imaging**, 2010. v. 23, n. 5, p. 611–631, 2010.
- RANGAYYAN, R. M.; PRAJNA, S.; AYRES, F. J.; DESAUTELS, J. E. L. Detection of architectural distortion in prior screening mammograms using gabor filters, phase portraits, fractal dimension, and texture analysis. **International Journal of Computer Assisted Radiology and Surgery**, 2008. v. 2, n. 6, p. 347–361, 2008.
- RAY, K. M.; TURNER, E.; SICKLES, E. A.; JOE, B. N. Suspicious Findings at Digital Breast Tomosynthesis Occult to Conventional Digital Mammography: Imaging Features and Pathology Findings. **The Breast Journal**, 2015. p. 1–5, 2015.
- RUSSO, P. **Handbook of X-ray Imaging: Physics and Technology**. [S.l.]: CRC Press, 2018. 1393 p. (Series in Medical Physics and Biomedical Engineering).
- SAIDIN, N.; SAKIM, H. A. M.; NGAH, U. K.; SHUAIB, I. L. Segmentation of Breast Regions in Mammogram Based on Density: A Review. **International Journal of Computer Science Issues**, 2012. v. 9, n. 4, p. 108–116, 2012.
- SCHINDELIN, J.; ARGANDA-CARRERAS, I.; FRISE, E.; KAYNIG, V.; LONGAIR, M.; PIETZSCH, T.; PREIBISCH, S.; RUEDEN, C.; SAALFELD, S.; SCHMID, B.; TINEVEZ, J.-Y.; WHITE, D. J.; HARTENSTEIN, V.; ELICEIRI, K.; TOMANCAK, P.; CARDONA, A. Fiji: an open-source platform for biological-image analysis. **Nature Methods**, 2012. Nature Publishing Group, v. 9, p. 676–682, 2012.
- SECHOPOULOS, I. A review of breast tomosynthesis. Part I. The image acquisition process. **Medical Physics**, 2013. v. 40, n. 1, 2013.

- SUCKLING, J.; PARKER, J.; DANCE, D.; ASTLEY, S.; HUTT, I.; BOGGIS, C.; RICKETTS, I.; STAMATAKIS, E.; CERNEAZ, N.; KOK, S. et al. The mammographic image analysis society digital mammogram database. In: **Exerpta Medica. International Congress Series**. [S.l.: s.n.], 1994. v. 1069, p. 375–378.
- SULEIMAN, W. I.; MCENTEE, M. F.; LEWIS, S. J.; RAWASHDEH, M. A.; GEORGIAN-SMITH, D.; HEARD, R.; TAPIA, K.; BRENNAN, P. C. In the Digital Era, Architectural Distortion Remains a Challenging Radiological Task. **Clinical Radiology**, 2016. The Royal College of Radiologists, v. 71, n. 1, p. e35–e40, 2016.
- TANG, J.; RANGAYYAN, R. M.; XU, J.; El Naqa, I. E.; YANG, Y. Computer-aided detection and diagnosis of breast cancer with mammography: Recent advances. **IEEE Transactions on Information Technology in Biomedicine**, 2009. v. 13, n. 2, p. 236–251, 2009.
- TINEVEZ, J.-Y.; PERRY, N.; SCHINDELIN, J.; HOOPES, G. M.; REYNOLDS, G. D.; LAPLANTINE, E.; BEDNAREK, S. Y.; SHORTE, S. L.; ELICEIRI, K. W. Trackmate: An open and extensible platform for single-particle tracking. **Methods**, 2017. v. 115, p. 80–90, 2017.
- VEDANTHAM, S.; KARELLAS, A.; VIJAYARAGHAVAN, G. R.; KOPANS, D. B. Digital Breast Tomosynthesis: State of Art. **Radiology**, 2015. v. 277, n. 3, p. 663–684, 2015.
- VERONESI, U.; BOYLE, P.; GOLDBIRSCH, A.; ORECCHIA, R.; VIALE, G. Breast Cancer. **The Lancet**, 2005. v. 365, p. 1727–1741, 2005.
- VIEIRA, R. T.; CHIERICI, C. E. de O.; FERRAZ, C. T.; GONZAGA, A. Local fuzzy pattern: A new way for micro-pattern analysis. In: YIN, H.; COSTA, A. F. J.; BARRETO, G. (Ed.). **Intelligent Data Engineering and Automated Learning - IDEAL 2012: 13th International Conference, Natal, Brazil, August 29-31, 2012. Proceedings**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. p. 602–611.
- WHO. World health organization - breast cancer. 2019. 2019. Disponível em: <<http://www.who.int/>>. Acesso em: 5 de abr. de 2019.
- WILSON, J.; RITTER, G. **Handbook of Computer Vision Algorithms in Image Algebra**. [S.l.]: CRC Press, 2000.
- YANG, S. K.; MOON, W. K.; CHO, N.; PARK, J. S.; CHA, J. H.; KIM, S. M.; KIM, S. J.; IM, J.-G. Screening mammography-detected cancers: Sensitivity of a computer-aided detection system applied to full-field digital mammograms. **Radiology**, 2007. v. 244, n. 1, p. 104–111, 2007.
- ZYOUT, I.; TOGNERI, R. Empirical Mode Decomposition of Digital Mammograms for the Statistical based Characterization of Architectural Distortion. In: **2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)**. [S.l.: s.n.], 2015. p. 109–112.
- ZYOUT, I.; TOGNERI, R. A computer-aided detection of the architectural distortion in digital mammograms using the fractal dimension measurements of BEMD. **Computerized Medical Imaging and Graphics**, 2018. v. 70, p. 173–184, 2018.