

**UNIVERSIDADE FEDERAL DO PIAUÍ – UFPI
CAMPUS SENADOR HELVÍDIO NUNES DE BARROS
CURSO DE BACHARELADO EM SISTEMAS DE INFORMAÇÃO**

**DESENVOLVIMENTO DE DESCRITORES BASEADO EM ANÁLISE DE
FORMA PARA DIAGNÓSTICO DE LESÕES PULMONARES**

ANTONINO CALISTO DOS SANTOS NETO

**PICOS – PIAUÍ
2017**

ANTONINO CALISTO DOS SANTOS NETO

**DESENVOLVIMENTO DE DESCRITORES BASEADO EM ANÁLISE DE FORMA
PARA DIAGNÓSTICO DE LESÕES**

Monografia submetida ao Curso de Bacharelado de Sistemas de Informação como requisito parcial para obtenção de grau de Bacharel em Sistemas de Informação.

Orientadora: Prof^a. Ma Alcilene Dalilia de Sousa

FICHA CATALOGRÁFICA
Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca José Albano de Macêdo

S237d Santos Neto, Antonino Calisto dos.

Desenvolvimento de descritores baseado em análise de forma para diagnóstico de lesões pulmonares / Antonino Calisto dos Santos Neto.– 2017.

CD-ROM : il.; 4 ¾ pol. (38f.)

Trabalho de Conclusão de Curso (Curso Bacharelado em Sistemas de Informação) – Universidade Federal do Piauí, Picos, 2017.

Orientador(A): Prof.^a Ma. Alcilene Dalília de Sousa

1. Câncer de Pulmão-Visão Computacional. 2.Descritores de Forma. 3. Imagens Médicas. I. Título.

CDD 005.118

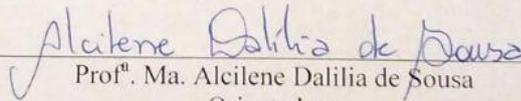
DESENVOLVIMENTO DE DESCRITORES BASEADO EM ANÁLISE DE FORMA
PARA DIAGNÓSTICO DE LESÕES PULMONARES

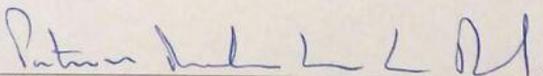
ANTONINO CALISTO DOS SANTOS NETO

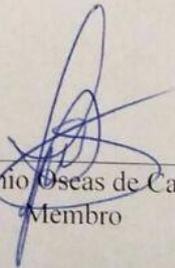
Monografia aprovada como exigência parcial para obtenção do grau de
Bacharel em Sistemas de Informação.

Data de Aprovação

Picos – PI, 19 de Janeiro de 2017


Prof.^a. Ma. Alcilene Dalília de Sousa
Orientadora


Prof.^a. Ma. Patrícia Medyna Lauritzen de Lucena Drumond
Membro


Prof. Dr. Antonio Oseas de Carvalho Filho
Membro

À Deus, e aos meus pais, Neusa Maria Calisto Alves e Jenival Alves da Silva por suas capacidades de acreditar em mim.

AGRADECIMENTOS

À Deus por iluminar meu caminho e me dar fôlego de vida para nunca desistir e seguir nessa longa jornada. Todas as conquistas e realizações só devo a ti.

À minha orientadora, Ma. Alcilene Dalília de Sousa, por me guiar desde o início da minha formação perante o estudo e a pesquisa. Obrigado pela positividade, paciência e ensinamentos transmitidos e pelo grande exemplo de determinação, lhe admiro muito quanto a isso, além da excelente coordenadora que é. Ao professor Dr. Antonio Oséas de Carvalho Filho, por transmitir sua vasta experiência no ramo da pesquisa, sempre à disposição por passar seu conhecimento, nos auxiliando a obter resultados significativos na pesquisa. E à professora Ma. Patrícia Medyna Lauritzen de Lucena Drumond, por ter paciência com os erros cometidos e por estar sempre à disposição com sua dedicação e competência.

Aos meus estimados professores, que no decorrer dessa caminhada, transmitiram valiosos conhecimentos. Em especial, agradeço ao Me. Romuere Rodrigues Veloso e Silva e ao Me. Flávio Henrique Duarte de Araújo, que mesmo distantes, se disponibilizavam a contribuir com seus conhecimentos no meu trabalho. Obrigado por toda ajuda que me deram!

Aos demais idealizadores, coordenadores e funcionários da Universidade Federal do Estado do Piauí - UFPI.

À minha mãe, Neusa Maria Calisto Alves, heroína que me deu apoio, guerreira de grande dedicação, amável e abençoada, que sempre me incentivou nas horas difíceis e sempre a seguir com humildade. À meu pai, que sempre me deu incentivo no estudo apesar das dificuldades, sempre estava disposto a ajudar, lutando sempre para dar o melhor aos seus filhos, muito obrigado.

Aos meus irmãos Jenival Alves da Silva Júnior e Paulyran Calisto Alves, que nas horas difíceis estavam dispostos a ajudar e a aconselhar.

À minha namorada, Carla do Vale Caminha, pela paciência, principalmente nesse último período de curso. Obrigado de coração por todo carinho, amor, respeito e por todo apoio nos momentos difíceis e por tornar minha vida cada dia mais feliz. Pessoa que admiro bastante, quanto ao comportamento pessoal e o modo de ver as coisas.

À minha família Calisto-Alves, que sempre me deu apoio para estudar e que

fizeram de tudo para que eu chegasse a essa etapa final do curso. De coração, muito obrigado.

Aos meus amigos do curso, Maura Gécica Rodrigues da Rocha, Marcos Raniere de Sousa Silva, Mirla Maria de Sousa Santana, Thayane de Oliveira Simões, Alexandre Ribeiro Cajazeira Ramos, Gabriel de Lima Leal, Tatiele Veloso da Silva, Gabriel Moura, Emerson Silva, Leonardo de Moura, e Luana Batista. Acredito que nossa amizade irá além da colação de grau.

A todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

"Determinação coragem e autoconfiança são fatores decisivos para o sucesso. Se estamos possuídos por uma inabalável determinação conseguiremos superá-los. Independentemente das circunstâncias, devemos ser sempre humildes, recatados e despidos de orgulho."

Dalai Lama

RESUMO

O câncer de pulmão é um dos tipos mais comuns de câncer e o que possui maior índice de mortalidade no mundo. É de grande importância o diagnóstico na fase inicial da doença, pois as chances de cura aumentam consideravelmente. A automatização do diagnóstico por sistemas de visão computacional, através da análise das imagens médicas, proporciona uma interpretação a respeito da patologia. A ideia é utilizar as características extraídas dos nódulos nas imagens, para que em seguida sejam classificados em nódulos malignos ou benignos, cujo resultado depende da qualidade das características extraídas. Este trabalho apresenta o desenvolvimento de descritores baseado na análise de forma para diagnóstico de lesões pulmonares. Foram obtidos resultados promissores, sendo um valor de especificidade de 86,2%, 91,5% de sensibilidade, 92% de Acurácia e uma curva ROC de 0.920.

Palavras-chave: Câncer de pulmão; diagnóstico; descritores de forma.

LISTA DE ILUSTRAÇÕES

Figura 1 - Comparação de dois objetos a partir do índice da compacidade: onde (A) é compacto e (B) é o não-compacto.....	20
Figura 2 - Demonstração de Feret.....	21
Figura 3 - Ilustração do retângulo do esqueleto, onde “A” e “B” pertencem ao esqueleto, e “C” não faz parte do mesmo.....	22
Figura 4. Ilustração do índice de densidade esférica em duas dimensões.....	24
Figura 5. TC de tórax em janela pulmonar. A: Contornos regulares e margens lisas. B: Contornos irregulares.....	25
Figura 6. Comparação de dois objetos a partir do índice da circularidade.....	26
Figura 7 - Método proposto.....	31

LISTA DE TABELAS

Tabela 1 - Descritores de forma utilizados no trabalhos de Sampaio (2015).....	23
Tabela 2 - Níveis de exatidão do índice <i>kappa</i>	28
Tabela 3 - Resultados para os múltiplos classificadores.....	33
Tabela 4 – Comparação entre os trabalhos relacionados e a metodologia proposta.....	34

LISTA DE ABREVIATURAS E SIGLAS

CADx	<i>Computer-Aided Diagnosis</i>
XML	eXtensible Markup Language
IDRI	Image Database Resource Initiative
INCA	Instituto Nacional do Câncer
LIDC	Lung Image Database Consortium
LIDC-IRDI	Lung Image Database Consortium
MAT	Medial Axis Transform
OSM	Organização Mundial da Saúde
PDI	Processamento Digital de Imagens
ROC	Receiver Operating Characteristic
SVM	Support Virtual Machine
TC	Tomografia Computadorizada
2D-DCT	Two Dimensional Discrete Cosine Transform
2D-FFT	Two Dimensional Fast Fourier Transform

SUMÁRIO

1 INTRODUÇÃO.....	15
1.1 OBJETIVOS.....	15
1.2 ORGANIZAÇÃO DO TRABALHO.....	16
2 ESTADO DA ARTE	17
2.1 CÂNCER DE PULMÃO E DIAGNÓSTICO	17
2.2 <i>PROCESSAMENTO DE IMAGENS</i>	18
2.3 <i>DESCRITORES</i>	19
2.4 <i>BASE DE IMAGENS</i>	19
2.5 <i>WEKA</i>	19
2.6 DESPROPORÇÃO ESFÉRICA.....	19
2.7 COMPACIDADE.....	20
2.8 MEDIDAS ESTADÍSTICAS	21
2.9 DESCRITORES ESTUDADOS EM SAMPAIO (2015)	21
2.9.1 Diâmetros de Feret.....	22
2.9.2 Esqueleto	22
2.10 ESFERICIDADE	24
2.11 DENSIDADE ESFÉRICA.....	25
2.12 IRREGULARIDADE	26
2.13 <i>CIRCULARIDADE</i>	27
2.14 <i>CLASSIFICAÇÃO</i>	28
2.15 <i>MÉTRICAS DE VALIDAÇÃO</i>	28
2.16 <i>TRABALHOS RELACIONADOS</i>	30
3 METODOLOGIA.....	32
3.3 AQUISIÇÃO DE IMAGENS.....	32
3.4 EXTRAÇÃO DE CARACTERÍSTICAS.....	33

3.5 CLASSIFICAÇÃO.....	33
4 RESULTADOS E DISCUSSÕES	34
5 CONSIDERAÇÕES FINAIS	36
REFERÊNCIAS BIBLIOGRÁFICAS.....	37

1 INTRODUÇÃO

Com o avanço da tecnologia, diversas áreas têm sido beneficiadas na simplificação de soluções para seus problemas. Uma dessas áreas é a área da saúde, destacando as pesquisas relacionadas ao diagnóstico do câncer.

Atualmente, o câncer de pulmão, é o câncer com maior taxa de mortalidade mundial no homem e o segundo na mulher, ficando atrás apenas para o câncer de mama (IARC,2015).

Com o intuito de aumentar a precisão na detecção do nódulos pulmonares surgiu o exame de Tomografia Computadorizada (TC). A TC surgiu como uma maneira mais eficaz para auxiliar na detecção de nódulos pulmonares, tornando-se bastante usual para o diagnóstico do câncer. Diferentes métodos foram desenvolvidos para auxiliar os médicos (e profissionais da área), no diagnóstico do câncer pulmonar, como os sistemas de Diagnóstico Auxiliado por Computador (“*Computer-Aided Diagnosis*” - CADx), cuja finalidade é melhorar a acurácia do diagnóstico (maligno e benigno), servindo de uma “segunda opinião” para a tomada de decisão.

A análise de um exame pode ocasionar alguns problemas no diagnóstico, por ser realizada de forma subjetiva, além da complexidade dos diferentes formatos que um nódulo pode assumir, ocasionando erros em sua análise, causada por fadiga visual, baixa qualidade de imagem, distração, entre outros. Faz-se necessário o desenvolvimento de novas maneiras de detectar e diagnosticar os nódulos pulmonares, com maior precisão.

Um descritor é um valor que descreve um determinado recurso de forma ou textura, de modo que seus resultados concordem com a intuição do especialista.

1.1 OBJETIVO

Este trabalho tem como objetivo apresentar o desenvolvimento de um descritor usando características de forma para classificação dos nódulos pulmonares em benignos ou malignos.

1.2 ORGANIZAÇÃO DO TRABALHO

O trabalho em questão está organizado em 5 (cinco) capítulos, que estão organizados da seguinte forma:

- Capítulo 2 – Estado da Arte: Nessa seção são apresentados os conceitos relacionados ao câncer de pulmão e o processamento de imagens, além da apresentação de trabalhos da literatura, utilizados como base de sustentação teórica da pesquisa.
- Capítulo 3 – Metodologia: São descritos e apresentados os materiais e métodos utilizados na pesquisa.
- Capítulo 4 – Resultados e Discussões: São apresentados e discutidos os resultados obtidos de acordo com os índices aplicados sobre a base de imagens utilizada no trabalho.
- Capítulo 5 – Conclusão e Trabalhos Futuros: São apresentadas as considerações finais da pesquisa desse trabalho, indicações de trabalhos futuros para possíveis testes de verificação de aperfeiçoamento.

2 ESTADO DA ARTE

Com a intenção de obter uma concepção mais clara sobre o problema proposto pelo presente estudo, fez-se necessária a formulação de conceitos e elaboração de ideias a respeito do tema, com o intuito de expor os pressupostos que darão suporte ao mesmo. Para isso, será disposta uma exposição conceitual sobre os principais itens que circundam o Diagnóstico Automático de Nódulos pulmonares utilizando Descritores de Forma e Múltiplos Classificadores.

2.1 CÂNCER DE PULMÃO E DIAGNÓSTICO

Segundo o Instituto Nacional do Câncer (INCA, 2016), o câncer é denominação para um grupo de mais de 100 doenças, caracterizado pelo crescimento desordenado das células, que invadem tecidos e células vizinhas, tornando-se, às vezes, muito agressivas e incontroláveis.

O câncer de pulmão é a principal causa de morte por câncer e está entre os tumores de maior incidência. Nos Estados unidos, o câncer de pulmão é o líder causador de mortalidade entre os tumores em ambos os sexos, representando 25% do total das morte por câncer e 5% do total das mortes em geral (SKARIN; BLANCO, 2013).

O câncer de pulmão é uma das neoplasias com a menor taxa de cura pelo fato de haver dificuldades no seu diagnóstico precoce. A sua classificação histológica mais aceita pela literatura mundial é a da Organização Mundial de Saúde (OSM), sendo os quatro tipos principais: Carcinoma espinocelular, adenocarcinoma, carcinoma indiferenciado de pequenas células e carcinoma indiferenciado de grandes células (UEHARA, JAMNIK; SANTORIO, 2015).

Quando o diagnóstico do câncer de pulmão é feito o mais cedo possível, maiores são as chances de cura. Os diagnósticos podem ser imprecisos por diversos fatores, como imagens ruins dos nódulos, grande quantidade de imagens para serem analisadas, fadiga visual por parte do radiologista, além de outras situações. O diagnóstico automático vem sendo desenvolvido para auxiliar os profissionais da área, extraíndo características dos nódulos nas imagens a partir da localização de seus padrões anormais. Análogo a essas ideias, pode-se definir o diagnóstico de

nódulos pulmonares, mais especificadamente, em malignos e benignos, sendo possível classificá-los por meio da análise de sua forma.

Um nódulo é caracterizado como uma pequena massa de tecido que se forma dentro ou sobre o corpo, geralmente devido a lesão. Os nódulos podem ser benignos, não sendo necessária prescrição médica, ou podem ser malignos, sendo necessário um auxílio médico em seu tratamento (SOUSA, 2007).

Os nódulos pulmonares, a maioria surgem nas paredes dos brônquios, atribuindo ao câncer de pulmão também o nome de bronco gênico, podendo também ocorrer nas paredes dos pulmões, levando, as vezes, vários anos para se desenvolverem.

2.2 PROCESSAMENTO DE IMAGENS

Uma imagem pode ser conceituada com uma função bidimensional $f(x,y)$, onde f é a informação da cor naquele específico ponto e x e y são as coordenadas espaciais. As técnicas de processamento de imagens, atualmente, estão sendo exploradas nas mais diversas áreas como, por exemplo, na radiologia, previsão de tempo, exploração espacial, radiologia e até nanotecnologia (RUSS, 2011).

Segundo mencionado por Jensen (2000), o principal objetivo do processamento digital de imagens, almeja a detecção, identificação e mensuração de objetos de uma cena, a partir de atividades multi-conceitos (como multiespectral, multitemporal e multiescalar) que abrangem princípios estatísticos para o reconhecimento de padrões, modelagem, sistemas de redes neurais, fotogrametria e interpretação de imagens.

As características visuais das imagens fornecem descrições de seu conteúdo. Baseando-se nesse conceito, pode-se utilizar descritores de formas para analisar o conteúdo contido nessas imagens.

Pelo fato do Processamento Digital de Imagens (PDI) ser a manipulação de imagens a partir de computadores, onde a imagem e a saída do processo são imagens, é possível fornecer técnicas que facilitem a identificação e a extração de características dessas imagens, e posteriormente realizar a classificação. Sendo assim, a partir da utilização de algoritmos de PDI, tornando-se possível a aplicação de várias técnicas para o diagnóstico automático do câncer de pulmão.

2.3 DESCRITORES

Um descritor de forma é um conjunto numérico que são produzidos para descrever um determinado recurso de forma, tentando quantificar a forma a ter uma interpretação condizente com a intuição humana (YANG *et al.*,2009). Muitos trabalhos utilizam esses descritores para extração de características, sendo bastante eficaz na descrição dos contornos ou formas da aplicação.

2.4 BASE DE IMAGENS

Para aplicação dos descritores desenvolvidos foi utilizada a base de imagens pública LIDC-IDRI, disponibilizada pelo *National Cancer Institute of EUA* (NCI). Esse banco de imagens foi criado a partir da combinação das bases *Lung Image Database Consortium (LIDC)* e o *Image Database Resource Initiative (IDRI)*. (ARMATO III *et al.*, 2011).

Essa junção resultou em uma base composta por 1018 exames, de acordo com as informações e marcações da base LIDC. Cada exame dessa base de imagens possui um arquivo XML (*eXtensible Markup Language*), com marcações das coordenadas dos nódulos e uma quantidade variável de fatias descritas por 4 especialistas (ARMATO III *et al.*,2011).

2.5 WEKA

A ferramenta utilizada para a classificação foi a WEKA (WEKA, 2016). Essa ferramenta contém pré-processamento de dados, classificação, regressão, clustering, regras de associação e visualização, sendo possível sua utilização para fazer o diagnóstico a partir das características a ele submetido, avaliando-as, que após um treinamento prévio, informa se o nódulos é de natureza maligna ou benigna.

2.6 DESPROPORÇÃO ESFÉRICA

Nódulos pulmonares benignos e maligno possuem características distintas principalmente quando analisados em três dimensões, podendo esses apresentarem aspectos mais arredondados (nódulos pulmonares benignos) ou aspectos com formas

mais alongadas (nódulos pulmonares malignos). Através da desproporção esférica é possível medir a regularidade desses aspectos.

A desproporção esférica (*DespEsf*) é medida por meio da comparação entre a área de um objeto com a área que este objeto ocuparia se fosse totalmente esférico, podendo assim, obter uma diferenciação quantitativa de sua estrutura morfológica (SOUSA, 2007).

A Equação 1 define como é mensurada a desproporção esférica de um objeto, onde A é a área da superfície do objeto, R é o raio estimado da esfera como o mesmo volume do objeto. O raio é obtido através da Equação 2, onde V é o volume do objeto.

$$DespEsf = \frac{A}{4\pi R^2} \quad (1)$$

$$R = \frac{\sqrt[3]{3V}}{4\pi} \quad (2)$$

Segundo Sousa (2007) o índice de desproporção esférica obtém valores maiores para objetos alongados ou irregulares e valores menores para objetos mais esféricos.

2.7 COMPACIDADE

Essa medida geométrica é calculada através da medição da densidade em relação a uma figura perfeitamente densa, ou seja, um círculo.

A Equação 3 apresenta a fórmula para calcular a compacidade (SAMPAIO, 2015), onde p representa área da superfície do objeto e V representa o volume interno do objeto.

$$C = \frac{V}{\frac{4\pi R^3}{3}} \quad (3)$$

Através da Figura 2 é possível observar graficamente o conceito de compacidade.

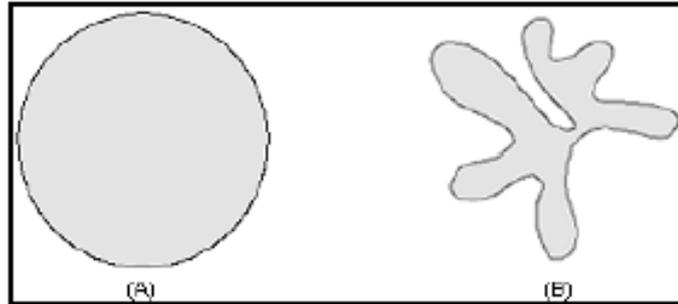


Figura 1 - Comparação de dois objetos a partir do índice da compactidade: onde (A) é compacto e (B) é o não-compacto.

Fonte: (SOUSA, 2007).

2.8 MEDIDAS ESTADÍSTICAS

Tendo em vista essa variação e com o intuito de produzir informações mais consistentes quanto ao diagnóstico, foi calculado o desvio padrão e a média da variância da distância entre o centro de massa do nódulo e a sua borda.

Esses cálculos podem ser observados nas Equações 4 e 5, onde M representa a média da distância entre o centro do candidato a nódulo e a sua borda, sendo n a quantidade de pontos identificados no decorrer de toda a borda. X_i indica a distância entre o centro do candidato a nódulo e um ponto i da borda.

$$M = \frac{\sum_i^n X_i}{n} \quad (4)$$

$$DP = \sqrt{\frac{\sum_i^n (X_i - M)^2}{n}} \quad (5)$$

2.9 DESCRITORES ESTUDADOS EM SAMPAIO (2015)

No trabalho de Sampaio (2015), foi utilizado atributos extras para posteriormente serem combinados e utilizados em descritores estudados por ele. As propriedades extras calculadas foram os Diâmetros de Feret e o tamanho do esqueleto.

2.9.1 Diâmetros de Feret

O diâmetro de Feret pode ser considerado a distância entre dois planos que definem o limite de um objeto por meio de retas perpendiculares a esse objeto e em uma direção específica. Essa técnica pode ser demonstrada na Figura 3.

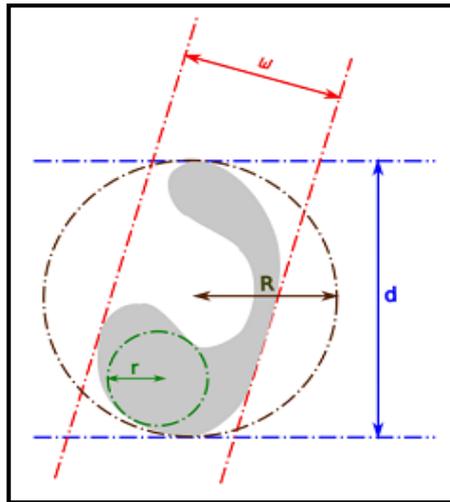


Figura 2 - Demonstração de Feret.

Fonte: Adaptado de (SAMPAIO, 2015).

Por meio da aplicação dessa técnica sobre uma imagem 3D, é possível extrair informações como volume do objeto (A), área da superfície (P), raio maior (R) e menor (r) de um circunferência interna da forma, e os diâmetros mínimo (w) e máximo (d) de Feret.

2.9.2 Esqueleto

Os descritores que utilizam o esqueleto como base, dividem-se em dois grupos. O primeiro grupo, é responsável pela obtenção as propriedades do esqueleto, e utiliza-las para extrair informações importantes a respeito da forma. O segundo grupo, é responsável por transformar o esqueleto em um grafo, e posteriormente é utilizado um algoritmo para casá-los (Sebastian, 2001). Pode-se citar, como exemplo relacionado ao primeiro grupo, as Saliências do Contorno, cujo conceito associa-se aos pontos de mais alta curvatura, combinando regiões de influência e onde esses contornos se localizam. Como exemplo para o segundo grupo, pode-se citar um descritor proposto

por Klein *et al.*, (2001), em que comparam grafos a partir da utilização de um algoritmo de distância de edição.

No algoritmo de distância de edição possui pontos, cujos são considerados pertencentes ao esqueleto, se ele for o centro de um círculo máximo e se este círculo tocar a borda em pelo menos dois pontos diferentes (SAMPAIO, 2015). Essa técnica foi utilizada em conjunto, através do tamanho do esqueleto (E_q) com os diâmetros mínimo (w) e máximo (d) de Feret para criar novos descritores, como pode ser observado nos índices 13 e 14 da Tabela 1.

Na Figura 4 é ilustrado a imagem do retângulo do esqueleto e a definição de seus centros e limites.

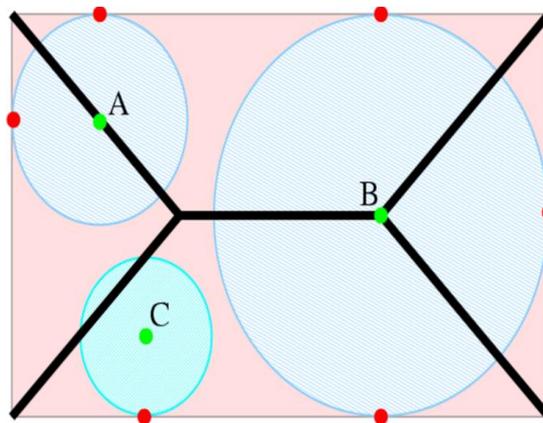


Figura 3 - Ilustração do retângulo do esqueleto, onde A e B pertencem ao esqueleto, e C não faz parte do mesmo.

Fonte: Adaptado de (SAMPAIO, 2015).

Com o intuito de uma melhor precisão das informações obtidas através da forma nódulos pulmonares, foram implementados alguns descritores de forma descritos no trabalho de Sampaio (2015), que levam em consideração as propriedades esféricas do nódulo, como comprimento, volume, raio máximo e mínimo da circunferência interna, entre outros. Estas propriedades esféricas foram obtidas através do estudo das propriedades dos nódulos feito por Sampaio (2015) e combinados com os outros descritores desse trabalho, com o intuito de uma extração com maior precisão.

Esses descritores podem ser observados na Tabela I, onde A é o volume do objeto, P é a área da superfície, R é o raio maior e r é o raio menor de uma circunferência interna da forma, w é o diâmetro mínimo e d é o diâmetro máximo de

Feret, e E_f é o tamanho do esqueleto.

Tabela 1 - Descritores de forma utilizados no trabalhos de Sampaio (2015).

Descritor	Fórmula	Descritor	Fórmula
1	$\frac{r}{R}$	8	$\frac{\pi w}{P}$
2	$\frac{w}{2R}$	9	$\frac{4\pi A}{P^2}$
3	$\frac{A}{2\pi R}$	10	$\frac{2d}{P}$
4	$\frac{2r}{d}$	11	$\frac{4R}{P}$
5	$\frac{4A}{\pi d^2}$	12	$\frac{2r}{w}$
6	$\frac{R\sqrt{3}}{d}$	13	$\frac{d}{E_f}$
7	$\frac{2\pi r}{P}$	14	$\frac{w}{E_f}$

Fonte: SAMPAIO (2015).

2.10 ESFERICIDADE

O cálculo da taxa de esfericidade (E) serve para medir o comportamento do nódulo, levando em consideração um objeto esférico (SILVA, 2007). Assim, a esfericidade (E) tem como objetivo medir o quão “perfeito” seria um nódulo, já que um nódulo perfeito, seria equivalente a uma esfera “perfeita” (SILVA, 2004). Aproxima-se de um formato esférico. Sua equação é definida pela fórmula abaixo:

$$E = \frac{6V\sqrt{\pi}}{A^{\frac{3}{2}}} \quad (6)$$

onde V é o volume do nódulo do nódulo, e A é a área da superfície do nódulo. Dessa forma, quanto mais próximo a forma do nódulo for da forma da esfera, mais o resultado se encontrará perto do valor 1 (SILVA, 2007).

2.11 DENSIDADE ESFÉRICA

Algumas medidas geométricas enfrentam problemas, quando utilizadas na comparação de uma área de um objeto com sua *bounding box*, ou bastante conhecida como caixa mínima. O *bounding box* é uma técnica em utiliza o menor cubo ou menos retângulo, voltados a objetos 3D, capaz de conter o objeto. O problema citado é devido essa medida possuir uma deficiência em calcular o alinhamento da figura circunscrita, com a menor distância das dimensões em relação ao objeto, de modo que sempre haja uma distância considerável entre a borda do objeto inscrito e o perímetro da figura, além de não ser uma técnica passível de rotações do objeto.

Tendo em vista esses problemas, pode-se utilizar a densidade esférica, que por ser uma esfera, possui características rotacionais invariantes. Sendo assim, com a densidade esférica é possível medir o quão compacto é um objeto, comparando o seu volume em relação ao volume de um *bounding box* (SOUSA; SILVA; PAIVA, 2008), os nódulos pulmonares possuem bastante semelhança com o formato de uma esfera, em contrapartida a outras estruturas como vasos, veias etc, que possuem pouca semelhança com formatos esféricos (SOUSA, 2007).

A fórmula da densidade esférica é definida por:

$$\text{DensEsf} = \frac{100 \cdot NV}{V} \quad (7)$$

Onde NV representa o número de voxels (pixel em 3D) do nódulo e V é o volume do nódulo.

A Figura 5 mostra graficamente o conceito de densidade esférica, em duas dimensões, para 3 objetos diferentes, sendo que de acordo com o círculo esférico no centro de cada objeto da imagem se é possível observar o primeiro objeto com um maior preenchimento, o segundo objeto com um parcial preenchimento em comparação com o primeiro e o ultimo objeto com pouco ou nenhum preenchimento, ilustrando assim, a dinamicidade do descritor proposto (Densidade Esférica).

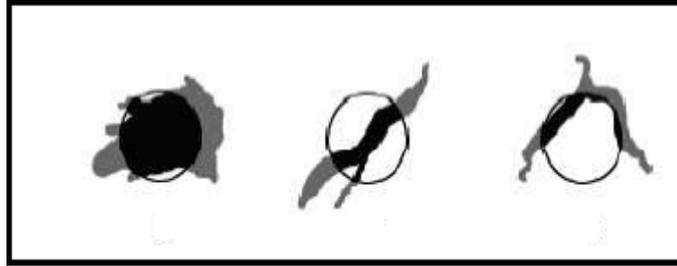


Figura 4 - Ilustração do índice de densidade esférica em duas dimensões.

Fonte: Adaptado de (KLEIN, 2001).

2.12 IRREGULARIDADE

A partir do momento que é definido o contorno de um nódulo, muitos atributos morfológicos podem ser determinados através de índices geométricos básicos. Um dos mais utilizados e comuns são o volume e a área.

Essas características são representadas graficamente e podem ser visualizadas na Figura 6.

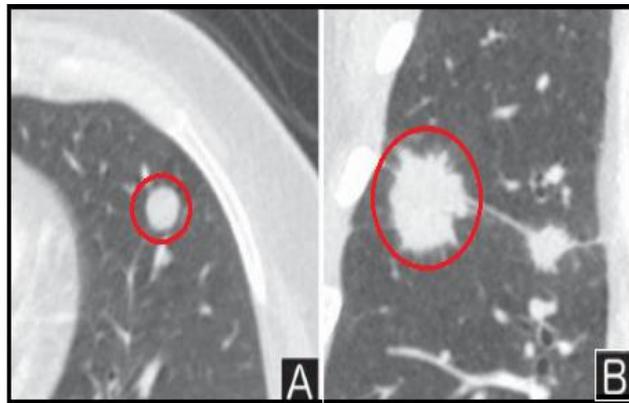


Figura 5 - TC de tórax em janela pulmonar. A: Contornos regulares e margens lisas. B: Contornos irregulares.

Fonte: Adaptado de (MOSMANN *et al.*, 2016).

A irregularidade de um nódulo é dado pela razão entre o volume e o quadrado da área de um nódulo (FRONER, 2015).

Sua fórmula pode ser definida por:

$$I = \frac{A^2}{V} \quad (8)$$

onde V representa o volume interno do nódulo e A representa a área da superfície do nódulo.

2.13 CIRCULARIDADE

A circularidade é caracterizada como uma medida geométrica que mede, através de um círculo, o quão próximo um objeto está (BRAZ JÚNIOR, 2014). O seu cálculo é feito através da razão entre o quadrado da área da superfície de um nódulo e seu volume.

A Equação 9 mostra como é calculado a circularidade.

$$C = \frac{A^2}{4\pi V} \quad (9)$$

onde A é a área da superfície do nódulo e V é o volume do nódulo. Assim, quanto mais próximo o objeto estiver de um círculo, mais próximo de 1 será o valor de sua circularidade (BRAZ JÚNIOR, 2014).

A circularidade é bastante eficaz na definição de formas, pois representa a medida específica do contorno dos nódulos que estão sendo estudados.

É possível observar um exemplo do conceito de circularidade na Figura 7, através da indicação de dois valores dessa medida para dois objetos.

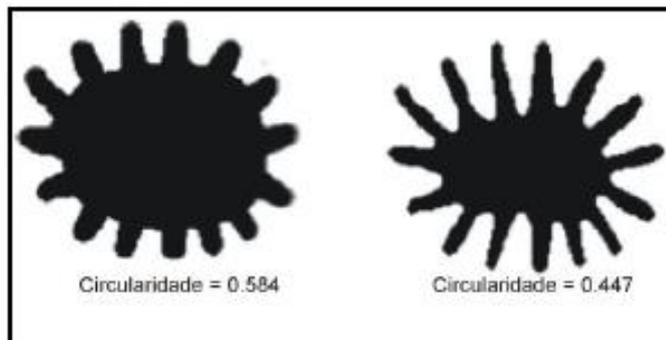


Figura 6 - Comparação de dois objetos a partir do índice da circularidade.

Fonte: (BRAZ JÚNIOR, 2014).

2.14 CLASSIFICAÇÃO

A classificação é um método de reconhecimento automático de objetos, onde possui entradas que são peculiaridades adquiridas através da etapa de extração sobre as imagens TC do pulmão.

Os métodos de Aprendizagem de Máquina são conceitos que se utilizam de técnicas para a extração de certos atributos de uma estrutura. Esse método utiliza-se da descoberta de elementos e a relação entre eles, com o intuito de chegar a uma possível conclusão (MANZANO-MANCHO; GÓMEZ-PÉREZ, 2005). Essa técnica é muito utilizada para o treinamento de algoritmos em processamento de imagens e posteriormente, após a extração dos atributos, sejam encaminhados para a etapa de classificação (GIGER, 1999).

2.15 MÉTRICAS DE VALIDAÇÃO

Esse trabalho usou-se de um teste de classificação que baseia-se em quatro possíveis ocorrências: Falso Positivo (FP), que é o número de imagens classificadas erradamente como não doentes; Verdadeiro Positivo (VP), que é quando os nódulos são classificados corretamente como doentes (maligno); Falso Negativo (FN), quando o número de imagens é classificado erradamente como não saudáveis e Verdadeiro Negativo (VN), quando o número de imagens é classificado corretamente como saudáveis (benigno).

As métricas utilizadas para validação dos resultados foram especificidade (ZHU; ZENG; WANG, 1998), sensibilidade (ZHU; ZENG; WANG, 1998), acurácia (PROVOST; DOMINGOS, 2000), índice Kappa (COOK, 1998) e a curva ROC (MARTINEZ; LOUZADA; PEREIRA, 2003).

A sensibilidade é conceituada como a probabilidade do teste em fornecer um resultado positivo, desde que o indivíduo seja portador da enfermidade. Ou seja, esse teste define o número de nódulos que são detectados corretamente (ZHU; ZENG; WANG, 1998).

A especificidade (*Esp*), por sua vez, se caracteriza como a probabilidade do fornecimento de um resultado negativo pelo teste. Ou seja, esse teste mede a proporção de pessoas sem a doença, indicando a qualidade do teste em identificar os

indivíduos não doentes, caracterizando o número de casos que foram identificados como não-nódulos (ZHU; ZENG; WANG, 1998).

Já acurácia, calcula o total de acertos baseado em todas as instâncias classificadas, calculado a partir de todos os casos (classificados corretamente e incorretamente) (PROVOST; DOMINGOS, 2000).

O índice *Kappa*, é considerada como uma das métrica mais importantes na etapa de classificação, representando uma totalidade na matriz de confusão. Ele mede a concordância formulada para medir o desacordo do conjunto de respostas baseando-se em observadores, para assim, chegar em um consenso (COOK, 1998).

A qualidade do índice *Kappa* é definida de acordo com o seu valor em uma escala de precisão, como pode ser observado na Tabela 2.

Tabela 2 - Níveis de exatidão do índice *kappa*.

Índide <i>Kappa</i> (K)	Qualidade
$K \leq 0,2$	Ruim
$0,2 < K \leq 0,4$	Razoável
$0,4 < K \leq 0,6$	Bom
$0,6 < K < 0,8$	Muito Bom
$K \geq 0,8$	Excelente

Fonte: Adaptado de (LANDIS; KOCH, 1977).

E por último, a curva ROC descreve, uma maneira aceitável e de grande utilização na literatura médica e na comunidade científica, de avaliar o desempenho quantitativo do método proposto, auxiliando na validação do teste (MARTINEZ; LOUZADA; PEREIRA, 2003).

A equação para cálculo da sensibilidade é definida pela razão entre número de VP e pelo total de nódulos existentes (dado por VP + FN), indicando assim a eficácia do teste. A equação pode ser observada na Equação 10.

$$S = \frac{VP}{VP+FN} \quad (10)$$

A equação para cálculo da especificidade (*Esp*) é calculado pela razão entre o

número de VN pelo número de casos que de fato não são nódulos (expresso pela soma de VN + FP). A equação pode ser observada na Equação 10.

$$Esp = \frac{VN}{VN+FP} \quad (11)$$

A equação para cálculo da Acurácia é expressa através razão do acerto total (dado por VP + VN) pelo total de casos da base. A equação para calculo da acurácia pode ser observado na Equação 10.

$$A = \frac{VP+VN}{VP+VN+FP+FN} \quad (12)$$

A partir da sensibilidade e da especificidade se é possível calcular a curva ROC para todas as observações da amostra. Define-se como melhor ponto da curva, o extremo superior esquerdo, representando à máxima sensibilidade e mínimo número de FPs, cujo resultado é considerado como de melhor performance, quando adquire um valor próximo a 1. Já o *Kappa*, representa uma medida de concordância entre observadores, que assim como a curva ROC, é caracterizado como de melhor performance, quando adquire um valor próximo de 1.

2.16 TRABALHOS RELACIONADOS

Na literatura, muitos trabalhos estão relacionados com o desenvolvimento de sistemas automáticos para diagnóstico do câncer de pulmão. Por esse motivo, são utilizadas características extraídas de imagens médicas, com o objetivo de classificar os nódulos em malignos ou benignos.

Na pesquisa Orozco *et al.* (2012), é apresentada uma alternativa computacional para classificar nódulos pulmonares na frequência em que se apresentam. As imagens são submetidas ao cálculo de dois espectros, o *Two Dimensional Discrete Cosine Transform* (2D-DCT) e o *Two Dimensional Fast Fourier Transform* (2D-FFT). Posteriormente, dois recursos estatísticos de textura foram extraídos do histograma e calculados a partir do espectro de cada imagem do CT. Em seguida, as características são passadas à SVM, onde 75 testes com

diagnósticos diferentes e o número de imagens foram usadas para validar a metodologia apresentada. Com isso, obtiveram o resultado de uma sensibilidade de 96,15, uma especificidade de 52,17% e uma acurácia de 82,66%.

Na pesquisa de Dandil *et al.*, (2014) apresentaram um método na qual utilizam imagens TC para diferenciação entre tumores malignos e benignos. O Sistema de CADx que foi desenvolvido forneceu uma segmentação de nódulos no qual se utilizou o *Self-Organizing Maps* e realizou a classificação entre nódulos benignos e malignos. Com isso, se obteve resultados com uma taxa de acerto de 90,63% de acurácia, uma taxa de acerto de 92,30% de sensibilidade e taxa de acerto de especificidade de 89,47%.

Orozco *et al.* (2015), propõe um método que avalia os descritores mais significativos utilizando uma *Support Virtual Machine* (SVM). A pesquisa conta com a validação de imagens de 45 TC das bases Lung Image Database Consortium – (LIDC-IRDI), onde obtiveram resultados promissores, com uma sensibilidade de 90,9%, Especificidade de 73,91% e acurácia de 82%.

Froner (2015) utilizou em seu trabalho atributos morfológicos, de intensidade e textura para uma melhor interpretação no diagnóstico auxiliado por computador. Ela utilizou uma avaliação para a utilização de dados de pacientes e atributos quantitativos de nódulos pulmonares em imagens de Tomografias Computadorizadas (TC) de pulmão para a elaboração de um modelo que melhor classificasse a malignidade dos nódulos, onde obteve um melhor resultado de uma provável curva ROC de 0.923 de precisão.

Diante do que foi encontrado na literatura, observou-se que as pesquisas citadas obtiveram uma aceitação considerável, com pouca variação entre sensibilidade, acurácia e especificidade. O trabalho proposto utilizou análise de forma para o diagnóstico automático de nódulos pulmonares, apresentado diversos descritores que podem ser utilizados para análise de forma, encontrando resultados promissores, com pequenos erros, devido a utilização de alguns descritores que não resultaram em uma extração promissora.

3 METODOLOGIA

A metodologia desenvolvida para a classificação dos nódulos pulmonares em maligno ou benigno foi empregada inicialmente com a aquisição das imagens; posteriormente, foram extraídas as características com a análise da forma desenvolvida a partir dos índices de forma; em seguida, utilizou-se a ferramenta WEKA para classificação em maligno e benigno, baseando-se na validação utilizando a sensibilidade, acurácia, especificidade, curva ROC e índice *Kappa* como está demonstrado na Figura 1.

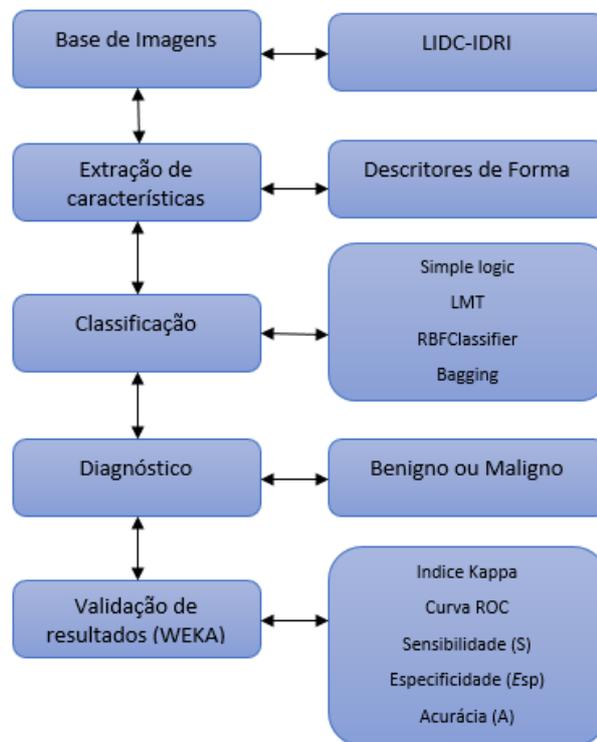


Figura 7 - Método proposto.

3.1 AQUISIÇÃO DE IMAGEM

Para aplicação dos algoritmos desenvolvidos, foram realizados testes de eficiência com a base LIDC-IDRI utilizando 1403 nódulos, sendo 1009 benignos e 394 malignos (ARMATO III *et al.*, 2011). Obteve-se como resultado essa detecção, devido essa base ser de grande complexidade e diversidade de exames, contando apenas

com exames convencionais de TC, mostrando a eficiência dos métodos propostos, onde obteve-se bons resultados se a exposição de uma maior radiação.

3.2 EXTRAÇÃO DE CARACTERÍSTICAS

A extração de característica é feita a partir das características da forma dos nódulos pulmonares como área, volume, esfericidade, entre outro, foi possível realizar a etapa de extração.

Os índices que foram estudados e utilizados com o intuito de obter um resultado que coincida com as opiniões dos especialistas, foram: Desproporção Esférica; Compacidade; Diâmetros de Feret; Esqueleto; Medidas estatísticas (medidas estatísticas relacionadas às distâncias da borda ao centro de Massa do Objeto); Descritores usados em Sampaio (2015); Esfericidade; Densidade esferica; Irregularidade; e Circularidade.

3.3 CLASSIFICAÇÃO

Os 4 classificadores aos quais esse trabalho foi submetido: *Simple logic* (LANDWEHR; HALL; FRANK, 2003), LMT (LANDWEHR; HALL; FRANK, 2003), *Bagging* (BREIMAN, 1996) e o *Radial Basis Function (RBF) Classifier* (FRANK, 2014) usaram o método *k-fold cross validation*, com $k = 10$. Os dados seguem uma divisão de 10 conjuntos, sendo 9 para treinamento e 1 para testes.

Utilizou-se diferentes classificadores com o intuito de verificar a eficiência do método em diferentes classificadores, estes escolhidos pelo fato de apresentarem o melhor resultado.

4 RESULTADOS E DISCUSSÕES

Esta seção refere-se aos resultados deste trabalho, que foi desenvolvido para auxiliar os especialistas no diagnóstico do cancer de pulmão. No decorrer do desenvolvimento, foram realizadas a revisão bibliográfica, o estudo do estado da arte, a implementação e testes, alguns algoritmos foram utilizados para extrair as características (chamados também de descritores) de forma, encontrados na literatura voltada à área de processamento de imagens.

Um teste de classificação deve ser avaliado, levando em consideração quatro perspectivas possíveis de ocorrências: FP, VP, FN e VN.

As métricas utilizadas para validação dos resultados foram especificidade (ZHU; ZENG; WANG, 1998), sensibilidade (ZHU; ZENG; WANG, 1998), acurácia (PROVOST; DOMINGOS, 2000), índice Kappa (COOK, 1998) e a curva ROC (MARTINEZ; LOUZADA; PEREIRA, 2003).

A Tabela 3 mostra que o classificador *SimpleLogistic* obteve o melhor resultado com um valor de 86,2% de especificidade (*E*), 91,5% de sensibilidade (*S*) e com um 92% de Acurácia (*A*), com um índice Kappa 0,6864 e uma curva ROC de 0,920. O erro do método proposto, é devido a grande variação das formas dos nódulos, havendo as vezes uma confusão entre malignos e benignos, por causa da semelhança das formas.

Tabela 3 - Resultados para os múltiplos classificadores.

Classificador	KAPPA	Curva ROC	S(%)	E(%)	A(%)
<i>Bagging</i>	0,6846	0,909	92,25	82	88
LMT	0,6864	0,920	91,5	85,2	91
<i>RBFClassifier</i>	0,6769	0,916	92,5	83,7	90
<i>SimpleLogistic</i>	0,6864	0,920	91,5	86,2	92

Como é possível observar na Tabela 3, a metodologia proposta apresenta uma pequena diferença do valor alcançado na primeira métrica (Sensibilidade), quando comparado ao equilíbrio entre as duas métricas (especificidade e acurácia), mas ainda capaz de indicar, de forma assertiva, que o paciente é portador da enfermidade.

Considerando os termos qualitativos do trabalho, a metodologia apresenta um resultado para o diagnóstico de nódulos pulmonares em malignos ou benignos, que condiz com as opiniões dos especialistas.

A Tabela 4 mostra alguns trabalho com resultados aceitos na literatura médica, comparados com a metodologia proposta.

Tabela 4 – Comparação entre os trabalhos relacionados e a metodologia proposta.

TRABALHOS	E(%)	A(%)	S(%)
(Orozco <i>et. al.</i> , 2015)	73,91	82	90,9
(Orozco <i>et. al.</i> , 2012)	52,17	82,66	96,15
(Dandil <i>et. al.</i> , 2014)	89,47	90,63	92,30
(Li, 2004)	75	93,7	80
Metodologia	86,2	92	91,5

De acordo com a Tabela 4, em comparação com os trabalhos relacionados, a metodologia proposta apresenta um resultado aceitável no que diz respeito às métricas. Assim, em termos qualitativos, a metodologia apresenta um bom resultado para o diagnóstico de nódulos pulmonares em maligno e benigno.

5 CONSIDERAÇÕES FINAIS

É notável que o diagnóstico de câncer de pulmão auxiliado por computador contribui de forma significativa para um diagnóstico correto e em menor tempo, pois representa um considerável aumento na probabilidade de sobrevivência dos pacientes.

Esse trabalho apresentou uma metodologia automática para classificação de nódulos pulmonares baseada na análise de forma com múltiplos classificadores com a intenção de diferenciar os nódulos pulmonares em maligno e benigno, servindo como uma segunda opinião para o especialista.

Os resultados obtidos confirmam um resultado promissor das técnicas de extração da forma, de acordo com a classificação na Tabela 2 e comparação com os trabalhos relacionados, com uma taxa de acerto de 92%, auxiliando ao especialista com uma segunda opinião no diagnóstico de nódulo pulmonar, proporcionando um tratamento precoce e com maiores chances de um prognóstico favorável ao paciente.

Para trabalhos futuros, pode-se implementar outros índices de formas, com a intenção de obter melhores resultados, pode-se também utilizar esses índices tanto em outras bases de imagens com TC, como também em bases de imagens de outros tipos de câncer.

REFERÊNCIAS BIBLIOGRÁFICAS

ARMATO III, S. G.; MCLENNAN G.; BIDAUT L.; MCNITT-GRAY M.F.; MEYER C.R.; REEVES A.P.; ZHAO B.; ABERLE D.R.; HENSCHKE C.I.; HOFFMAN E.A.; et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): a completed reference database of lung nodules on CT scans, **Med. Phys.**, v. 38, p. 915–931, 2011.

BLUM, H. **A transformation for extracting new descriptors of form In Models for the Perception of Speech and Visual Form, Whaten-Dunn (Ed.)** MIT Press: Cambridge, MA, p. 362-380, 1967.

BRAZ JÚNIOR, G.. **Detecção de regiões de massas em mamografias usando índices de diversidade geoestatística e geometria côncava.** Tese (Doutorado na Área de Ciência da Computação) - Programa de Pós-graduação em Ciências da Computação, Centro de Ciências e Tecnologia, Universidade Federal do Maranhão, São Luis, 2014.

BREIMAN, L.. Bagging Predictors, in:_____. **Machine Learning**, online, 1996, v. 24, n. 2, p. 123-140.

CHEN, W.; LI, Z.; BAI, L.; LIN, Y.. **NF- kappaB, a mediator for lung carcinogenesis and a target for lung cancer prevention and therapy,** *Frontiers in bioscience: a journal and virtual library.* v. 16, p. 1172-1185, 2011.

CONFERENCE ON EVOLUTIONARY PROGRAMMING. 1998, San Diego. **Sensitivity, specificity, accuracy associated confidence interval and roc analysis with practical sas implementations.** San Diego: Proceedings of the NorthEast SAS Users Group, CA, 1998, p. 201-208.

COOK, R. J.. **Kappa and its Dependence on Marginal Rates.** *Encyclopedia of BioStatistics.* P. Armitage and T. Colton, eds., John Wiley &, Sons, 1998, p. 2166--2168.

DANDIL, E.; CAKIROGLU, M.; EKSI, Z.; OZKAN, M.; KURT, O.; CANAN, A.. **Artificial neural network-based classification system for lung nodules on computed tomography scans.** *Soft Computing and Pattern Recognition (SoCPaR) 2014 6th International Conference of.* 2014. pp. 382-386.

FRANK, E.. **Fully supervised training of Gaussian radial basis function networks in WEKA.** Department of Computer Science, The University of Waikato, 2014.

FRAZÃO, A.. **Que tipo de nódulo pode ser câncer:** Tua saúde. Disponível em: <https://www.tuasaude.com/que-tipo-de-nodulo-no-pulmao-pode-ser-cancer/>>. Acesso em: 15 Nov. 2016.

GIGER, M. L.. **Computer-aided diagnosis.** *RSNA Categorical Course in Breast Imaging,* 1999, pp. 249-72.

INTERNATIONAL AGENCY FOR RESEARCH ON CANCER - IARC. **About CANCERModial**. Disponível em: <<http://www-dep.iarc.fr>>. Acesso em: 17 jun. 2016.

JENSEN, J. R. **Introductory digital image processing: a remote sensing perspective**. New Jersey: Prentice Hall, p. 379.1986.

KLEIN, P. N.; TIRTHAPURA, S.; SHARVIT, D.; KIMIA, B. B.. **A tree-edit-distance algorithm for comparing simple, closed shapes**. In Symposium on Discrete Algorithms, 2000, p. 696–704.

LANDIS, J. R.; KOCH, G. G.. **The measurement of observer agreement for categorical data**, Biometrics, 1977, v. 33, p.159-174.

LI F.; AOYAMA M.; SHIRAIISHI J.; ABE H.; LI Q.; SUZUKI K.; et al. Improvement in radiologists performance for differentiating small benign from malignant lung nodules on high-resolution CT by using computer-estimated likelihood of malignancy. **AJR Am J Roentgenol**, v. 183, n.5, p. 1209-1215, 2004.

MANZANO-MANCHO, D.; GÓMEZ-PÉREZ, A.. An overview of methods and tools for ontology learning from texts. **The Knowledge Engineering Review**, Set. 2004, vol 19, n.3, p. 187-212.

MARTINEZ, E. Z.; LOUZADA, N. F; PEREIRA, B. B.. **A curva ROC para tests diagnósticos**. Cadernos Saúde Coletiva 11, 2003, pp. 7-31.

METZ, C. E.. **ROC methodology in radiologic imaging**. Invest. Radiol. 1986. v. 21, n.9, p. 720-33.

MOSMANN, M. P.; BORBA, M. A.; MACEDO, F. P. N. D.; LIGUORI, A. D. A. L.; NETO, A. V.; LIMA, K. C. D.. Nódulo pulmonar solitário e 18F-FDG PET/CT. Parte 1: epidemiologia, avaliação morfológica e probabilidade de câncer, **Radiologia Brasileira**, V. 49, N. 1, p. 35-42, Jan/Fev. 2016.

NACIONAL CANCER INSTITUTE - NCI. **Lung Cancer. U.S. National Institute of Health**. Disponível em: <www.cancer.gov/about-cancer/what-is-cancer> Acesso em: 17 jun. 2016).

OROZCO, H. M. et al.. **Lung nodule classification in frequency domain using support vector machines**. in Information Science, Signal Processing and their Applications (ISSPA), 2012 11th International Conference on. 2012. pp. 870-875.

OROZCO, H.M.; VILLEGAS, O. O. V.; SANCHEZ, V. G. C.; DOMINGUEZ, H. D. J. O.; ALFARO, M. D. J. N. Automated system for lung nodules classification based on wavelet feature descriptor and support vector machine. **Biomedical Engineering Online**, v.14, n. 9, Jan. 2015.

PROCEEDINGS OF THE EUROPEAN CONFERENCE ON MACHINE LEARNING LANDWEHR, 14th, 2004, Croácia, **Anais**. 2006 Kluwer Academic Publishers: Printed in the Netherlands, v. 2837, p. 241–252.

PROVOST, F.; DOMINGOS, P.; **Well-Trained Pets: Improving Probability Estimation Trees**. Stern School of Business. Universidade de New York. 2000.

RUSS, J. C. **The image processing handbook**. 6^a ed. Florida: CRC Press. P. 885. 2011.

SAMPAIO, W. B.. **Detecção de massas em imagens mamográficas usando uma metodologia adaptada à densidade da mama**. Tese (Doutorado na Área de Engenharia Elétrica) – Programa de Pós-Graduação em Engenharia de Eletricidade, Centro de Ciências e Tecnologia, Universidade Federal do Maranhão, São Luis, 2015.

SEBASTIAN, T. B.; KLEIN, P. N.; KIMIA, B. B.. **Recognition of shapes by editing shock graphs**, in IEEE ICCV, 2001.

SILVA, A. C.. **Algoritmos para Diagnóstico Assistido de Nódulos Pulmonares Solitários em Imagens de Tomografia Computadorizada**. Tese (Doutorado na Área de Informática) - Programa de Pós-graduação em Informática, Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2004.

SILVA, A. C.. **Medidas globais em 3d para diagnóstico de nódulo pulmonar**. Disse, Instituto de Matemática Pura e Aplicada–IMPA, 2007.

SKARIN, A. T.; BLANCO, R.. **Lung cancer and tumors of the heart and mediastinum**. 2nd ed. Skarin AT: Atlas of diagnostic oncology. Philadelphia: Mosby-Wolfe. 2013. p. 63-109.

SOUSA, J. R. F. S. **Metodologia para detecção automática de nódulos pulmonares**. Dissertação (Mestrado na Área de Engenharia Elétrica) - Programa de Pós-graduação em Engenharia da Eletricidade, Centro de Ciências e Tecnologia, Universidade Federal do Maranhão, São Luis, 2007.

SOUSA, J. R. F. S.; SILVA, A. C.; PAIVA, A. C.. Lung structure classification using 3D geometric measurements and SVM. Progress in Pattern Recognition, Image Analysis and Applications.. **Lecture Notes in Computer Science**., v. 4756, p. 783-792, 2008.

UEHARA, C.; JAMNIK, S.; SANTORO, I. L.. **Câncer de pulmão**. Medicina (Ribeirao Preto. Online), v. 31, n. 2, p. 266-276, 2015.

WEKA – Machine Learning Group at the University of Waikato. **Weka 3: Data Mining Software in Java**. Disponível em: <<http://www.cs.waikato.ac.nz/ml/weka/>>. Acesso em: 12 Nov. 2016.



TERMO DE AUTORIZAÇÃO PARA PUBLICAÇÃO DIGITAL NA BIBLIOTECA “JOSÉ ALBANO DE MACEDO”

Identificação do Tipo de Documento

- () Tese
() Dissertação
(X) Monografia
() Artigo

Eu, Antonino Calisto dos Santos Neto, autorizo com base na Lei Federal nº 9.610 de 19 de Fevereiro de 1998 e na Lei nº 10.973 de 02 de dezembro de 2004, a biblioteca da Universidade Federal do Piauí a divulgar, gratuitamente, sem ressarcimento de direitos autorais, o texto integral da publicação Desenvolvimento de descritores baseado em análise de forma para diagnóstico de lesões pulmonares de minha autoria, em formato PDF, para fins de leitura e/ou impressão, pela internet a título de divulgação da produção científica gerada pela Universidade.

Picos-PI 07 de Fevereiro de 2017.

Antonino Calisto dos Santos Neto

Assinatura