UNIVERSIDADE DE SÃO PAULO
FACULDADE DE MEDICINA DE RIBEIRÃO PRETO

LUCAS FRIGHETTO PEREIRA

Classificação semiautomática de fraturas vertebrais benignas e malignas
em imagens de ressonância magnética

Ribeirão Preto
2016

LUCAS FRIGHETTO PEREIRA

Classificação semiautomática de fraturas vertebrais benignas e malignas em imagens de ressonância magnética

Dissertação apresentada à Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo para obtenção do título de Mestre em Ciências.

Orientador: Prof. Dr. Marcello Henrique Nogueira-Barbosa

Ribeirão Preto
2016

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Frighetto-Pereira, Lucas

Classificação semiautomática de fraturas vertebrais benignas e malignas em imagens de ressonância magnética, 2016.

129 p.: il.; 30 cm

Dissertação de Mestrado, apresentada à Faculdade de Medicina de Ribeirão Preto/USP.

Orientador: Nogueira-Barbosa, Marcello Henrique.

1. Diagnóstico auxiliado por computador.
2. Fratura vertebral por compressão.
3. Imagens de ressonância magnética.
4. Processamento de imagens.

FOLHA DE APROVAÇÃO

Lucas Frighetto Pereira

“Classificação semiautomática de fraturas vertebrais benignas e malignas
em imagens de ressonância magnética”

Dissertação de Mestrado apresentada a Faculdade de
Medicina de Ribeirão Preto da Universidade de São Paulo
para obtenção do título de Mestre em Ciências

Orientador: Prof. Dr. Marcello Henrique Nogueira-Barbosa

Aprovado em: ____/____/____

Banca Examinadora

Prof(a). Dr(a). _____

Instituição: _____ Assinatura: _____

Prof(a). Dr(a). _____

Instituição: _____ Assinatura: _____

Prof(a). Dr(a). _____

Instituição: _____ Assinatura: _____

DEDICATÓRIA

Dedico esse trabalho à minha **Família**,
que além de ser minha base de tudo e meu porto seguro,
foi onde recebi a melhor de todas as educações

Ao meu orientador, **Prof. Marcello**,
grande como profissional e enorme como ser humano
que me abriu grandes portas e esteve comigo nesse desenvolvimento profissional

À **FAPESP**,
que além de financiar meu mestrado,
me proporcionou experiências incríveis que levarei onde quer que eu vá

Aos meus **amigos do LAPIR-CCIFM**,
grandes amigos da carinhosamente chamada Borracharia,
que muito me ajudaram e me aturaram por todos esses anos

À **Cia do Riso**,
que me ensinou muito sobre o valor da vida
onde evolui muito como ser humano usando um nariz redondo e vermelho

À **Deus**,
o qual tenho muito a agradecer!

AGRADECIMENTOS

Agradeço primeiramente a Deus por ter nascido em uma família maravilhosa que me educou como ser humano e me permitiu estudar para chegar até aqui. Agradeço a Ele pelas portas abertas, pelas pessoas boas colocadas em meu caminho e, claro, por me dar a oportunidade de realizar meu mestrado em uma das melhores universidades do mundo, a Universidade de São Paulo.

Agradeço à minha família, em especial aos meus pais Ailton e Márcia e ao meu irmão Leonardo, que são minha base de tudo. Deram-me segurança, educação, bons conselhos, suporte, e motivação para vencer. Dedico não somente meu mestrado, mas dedico toda minha vida a eles.

Agradeço aos meus avós maternos Fernando e Laura e paternos Marcilio e Dirce. Espero que meus filhos e netos sintam por mim o que eu sinto por vocês. Um grande abraço aos meus tios Fernando, Eduardo, Giroto, Tia Cidinha, Tia Lu, Tio Mingo e à minha prima predileta Mariela.

Agradeço ao meu orientador Professor Marcello Henrique que tornou tudo isso possível. Excelente profissional e sensacional ser humano. Muito obrigado por todos os ensinamentos e oportunidades. Agradeço a você por tudo que fiz e conquistei durante meu período de mestrado: bolsa FAPESP, ao intercâmbio pela FAPESP (BEPE), ao PAE, ao Pequeno Cientista e todo aprendizado acadêmico adquirido.

Um agradecimento mais que especial aos professores Paulo Mazzoncini e Raj Rangayyan que foram fundamentais no desenvolvimento de meu trabalho e no meu aprendizado acadêmico. Agradeço muito ao Professor Rangayyan e ao meu amigo Faraz pela hospitalidade durante minha inesquecível experiência na Universidade de Calgary no Canadá.

Agradeço a FAPESP pelas minhas duas bolsas concedidas (2014/12135-0 e 2015/08778-6). Sensacional agência de fomento. Sem a FAPESP nada seria possível. Muito

obrigado por acreditar em meu projeto e me proporcionar essa experiência sensacional com a BEPE. Tenho muito orgulho em falar que fui bolsista da FAPESP. Muito obrigado!

Agradeço a turma da Borracharia, vulgo LAPIR do CCIFM do HC das Clínicas: Spiderman Rafael ~~dos Santos~~ de Menezes Reis, Júlio Cezar, Gui Modafoca Zeri, dotô Vitor Barriguinha, Dr. Cana, menino Hilton, menino Neto, ao Santiago Papae Baêa, Marcelo Benito da Beni TI, ~~Jaqueti~~ Jaqueline, Rodrigão, Yago do Amendoim, mestre e chefe Gustavo Bonúglius, Raniery, dotô Everaldo, Débora, nosso herói Goku e à Dona Elza pelo café de cada dia, Amém. Muitas histórias, risadas, muito trabalho, aprendizado e sensacionais experiências que vivi com vocês. Cada um de vocês contribuiu para essa minha conquista, com um especial ao menino Rafael que foi um irmão para mim durante todos esses anos.

Para finalizar, gostaria de dizer novamente: “muito obrigado CIA DO RISO”. Aprendi muito com vocês! Fiz grandes amigos durante quase sete anos que levarei para sempre em minha vida. Ganhei um irmão gêmeo palhaço chamado Paraguai que, com um jaleco e um nariz redondo e vermelho, me ensinou muito sobre o valor que a vida tem. Gostaria de terminar os agradecimentos e iniciar minha dissertação de mestrado com uma frase inesquecível que tive o prazer de ouvir de uma mãe de uma criança internada após uma visita em seu quarto:

“Palhaço, OBRIGADO PELA FELICIDADE”

LISTA DE FIGURAS

- Figura 1 - Ilustração de três IRM da coluna vertebral de diferentes pacientes em corte sagital mediano. Em (a) os corpos vertebrais estão íntegros não havendo presença de fraturas vertebrais. (b) apresenta duas FVCs nos corpos vertebrais L1 e L4 de características benignas. (c) apresenta duas FVC nos corpos vertebrais L1 e L5 com características de metástase óssea..... (3)
- Figura 2 - Deformidades dos corpos vertebrais em graus de lesões de acordo com Genant et al (1993)..... (6)
- Figura 3 - Histograma de níveis de cinza de uma ROI de 8-bits/pixel de um corpo vertebral com FVC benigna. Note que apenas a região do corpo vertebral foi utilizada para o cálculo do histograma, excluindo o fundo preto..... (10)
- Figura 4 - Imagem 5x5 em uma escala com três níveis de cinza: (a) representação visual; (b) representação computacional..... (11)
- Figura 5 - Em destaque está a matriz de coocorrência de níveis de cinza da imagem da Figura 4 com distância = 1 e ângulo = $\pm 45^\circ$ (12)
- Figura 6 - Ilustração da segmentação manual de um exame de um paciente com uma FVC benigna em L1. (a) Exame de RM em ponderação T1; (b) Máscara do exame de RM gerada pela segmentação manual..... (25)
- Figura 7 - Fluxograma do software de extração de características desenvolvido. O processo ocorre por 63 vezes, que é a quantidade de imagens selecionadas para o trabalho..... (26)
- Figura 8 - Imagem de máscara binarizada. Os corpos vertebrais estão representados em branco, os centroides são os pontos em vermelho e as *bounding boxes* estão representadas em verde..... (27)

- Figura 9 - Exame de RM com uma FVC maligna em L4 secundária a metástase de carcinoma mamário. A seta branca aponta para L4 e a seta amarela aponta o disco intervertebral inferior a L4. Note a semelhança de intensidade de sinal de ambos..... (28)
- Figura 10 - Os contornos dos corpos vertebrais estão mostrados nas linhas verdes. As *bounding boxes* de cada corpo vertebral estão mostradas nas linhas amarelas. Em azul são regiões dos discos intervertebrais gerados automaticamente utilizando as *bounding boxes* dos corpos vertebrais como referência. Em vermelho estão quatro blocos de 5x5 pixels cada, automaticamente posicionados nos discos intervertebrais entre L1-L2, L2-L3, L3-L4 e L4-L5 utilizando as *bounding boxes* vertebrais como referência..... (29)
- Figura 11 - Exemplo de imagem resultante da Etapa 4 do software..... (30)
- Figura 12 - Exemplo de duas ROIs e suas *bounding boxes* em amarelo. A seta branca indica o ruído do corpo vertebral adjacente envolvido pela *bounding box* que é removido pelo software..... (31)
- Figura 13 - *Bounding box* com ruído apresentada na Figura 12. (a) a seta branca aponta para o ruído na imagem; (b) a seta branca destaca que o ruído foi removido. (31)
- Figura 14 - Exemplos de imagens de ROIs dos corpos vertebrais geradas na Etapa 5 do software. Através dessas ROIs são extraídas as características dos corpos vertebrais. A linha (a) representa três exemplos de ROIs com corpos vertebrais com padrão de normalidade; a linha (b) representa três exemplos de ROIs com FVCs de padrão benigno; a linha (c) representa três exemplos de ROIs com FVCs de padrão maligno..... (32)
- Figura 15 - Mesmas ROIs dos corpos vertebrais ilustrados na Figura 14 após a binarização para a utilização nos cálculos de atributos de forma. A linha (a) representa três exemplos de ROIs com corpos vertebrais com padrão de normalidade; a linha (b) representa três exemplos de ROIs com FVCs de padrão benigno; a linha (c) representa três exemplos de ROIs com FVCs de padrão maligno..... (38)

- Figura 16 - Representação do contorno do corpo vertebral na cor vermelha onde o perímetro P é calculado e na cor branca onde a área vertebral V_{area} foi calculada..... (39)
- Figura 17 - O polígono em verde representa o *convex hull* do corpo vertebral e em vermelho está o contorno do corpo vertebral..... (40)
- Figura 18 - Em ciano estão ilustradas as distâncias entre o contorno dos corpos vertebral e seus centroides em vermelho. (a) corpo vertebral dentro da normalidade; (b) FVC benigna; (c) FVC maligna..... (44)
- Figura 19 - As linhas em vermelho representam o eixo principal das ROIs, as linhas azuis perpendiculares ao eixo principal representam as alturas superior-inferior e as linhas verdes paralelas ao eixo principal representam as larguras anterior-posterior. (a) ROI de um corpo vertebral normal, (b) ROI de um corpo vertebral com uma FVC padrão benigno e (c) ROI de um corpo vertebral com uma FVC padrão maligno..... (45)
- Figura 20 - Ilustração da classificação de corpos vertebrais com FVC x Normal..... (48)
- Figura 21 - Ilustração da classificação de corpos vertebrais com FVC Maligna x FVC Benigna..... (49)
- Figura 22 - Ilustração da classificação de corpos vertebrais com FVC Maligna x FVC Benigna x Normal..... (49)
- Figura 23 - Fluxograma da seleção e classificação dos vetores de características: a representa o vetor de características com todos os atributos que é utilizado diretamente para a classificação em todos os classificadores; b representa o vetor de características com todos os atributos que é submetido ao método automático Wrapper de seleção de atributos para cada classificador em específico resultando em $b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8$; c representa o vetor de características com todos os atributos que é submetido ao método seleção de

atributos 5+ A_z resultando em c_1 que é utilizado para a classificação em todos os classificadores; d representa o vetor de características com todos os atributos que é submetido ao método seleção de atributos 10+ A_z resultando em d_1 que é utilizado para a classificação em todos os classificadores..... (52)

LISTA DE TABELAS

Tabela 1 -	Exemplo de uma matriz de confusão	(14)
Tabela 1 -	Base de dados do trabalho quantificando os corpos vertebrais lombares de acordo com a classificação padrão-ouro.....	(24)
Tabela 2 -	Atributos estatísticos calculados a partir das alturas superior-inferior, das larguras anterior-posterior e da união de ambas.....	(45)
Tabela 3 -	Valores de p-valores calculados pelo teste Wilcoxon-Mann-Whitney e os valores de A_z para cada atributo na análise FVC x Normal. Cores das condições: $p < 0,001$ e $A_z \geq 0,9$ (preenchimento cinza escuro e escrito em branco); $0,001 \leq p \leq 0,01$ e $0,8 \leq A_z \leq 0,89$ (preenchimento cinza intermediário e escrito em preto); $0,01 \leq p \leq 0,05$ e $0,7 \leq A_z \leq 0,79$ (preenchimento cinza claro e escrito em preto); $p > 0,05$ e $A_z < 0,7$ ou valor desconhecido (?) (sem preenchimento de cor e escrito em preto).....	(57)
Tabela 4 -	Valores de p-valores calculados pelo teste Wilcoxon-Mann-Whitney e os valores de A_z para cada atributo na análise FVC Maligna x FVC Benigna. Cores das condições: $p < 0,001$ e $A_z \geq 0,9$ (preenchimento cinza escuro e escrito em branco); $0,001 \leq p \leq 0,01$ e $0,8 \leq A_z \leq 0,89$ (preenchimento cinza intermediário e escrito em preto); $0,01 \leq p \leq 0,05$ e $0,7 \leq A_z \leq 0,79$ (preenchimento cinza claro e escrito em preto); $p > 0,05$ e $A_z < 0,7$ ou valor desconhecido (?) (sem preenchimento e escrito em preto).....	(59)
Tabela 5 -	Cinco atributos com maior valor de A_z selecionados para compor o vetor de características dos corpos vertebrais na classificação de FVC x Normal.....	(61)
Tabela 6 -	Cinco atributos com maior valor de A_z selecionados para compor o vetor de características dos corpos vertebrais na classificação de FVC Maligna x FVC Benigna.....	(62)
Tabela 7 -	Dez atributos com maior valor de A_z selecionados para compor o vetor de características dos corpos vertebrais na classificação de FVC x Normal.....	(63)

- Tabela 8 - Dez atributos com maior valor de A_z selecionados para compor o vetor de características dos corpos vertebrais na classificação de FVC Maligna x FVC Benigna..... (63)
- Tabela 9 - Seleções de atributos pelo método *Wrapper*, com *cross-validation* de *folds* = 10, método de busca *best first* para cada classificador e para a análise “FVC x Normal”. Os atributos selecionados 30% dos *folds* ou mais são os atributos que irão compor os vetores características do respectivo classificador e estão destacados em negrito. Em (?) são valores não obtidos..... (64)
- Tabela 10 - Seleções de atributos pelo método *Wrapper*, com *cross-validation* de *folds* = 10, método de busca *best first* para cada classificador e para a análise “FVC Maligna x FVC Benigna”. Os atributos selecionados 30% dos *folds* ou mais são os que irão compor os vetores características do respectivo classificador e estão destacados em negrito. Em (?) são valores não obtidos..... (65)
- Tabela 11 - Seleções de atributos pelo método *Wrapper*, com *cross-validation* de *folds* = 10, método de busca *best first* para cada classificador e para a análise “FVC Maligna x FVC Benigna x Normal”. Os atributos selecionados 30% dos *folds* ou mais são os que irão compor os vetores características do respectivo classificador e estão destacados em negrito. Em (?) são valores não obtidos..... (66)
- Tabela 12 - Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com *cross-validation* de 10-*folds*, utilizando o vetor de características com todos os atributos calculados para os corpos vertebrais para a análise “FVC x Normal”. Em negrito estão os valores das médias aritméticas dos resultados das classificações..... (67)
- Tabela 13 - Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com *cross-validation* de 10-*folds*, utilizando o vetor de características com todos os atributos calculados para os corpos vertebrais para a análise “FVC Maligna x FVC Benigna”. Em negrito estão os valores das médias aritméticas dos resultados das classificações..... (67)

- Tabela 14 - Acurácia, AUROC de cada classificação, com *cross-validation* de 10-folds, utilizando o vetor de características com todos os atributos calculados para os corpos vertebrais para a análise “FVC Maligna x FVC Benigna x Normal”. Em negrito estão os valores das médias aritméticas dos resultados das classificações..... (68)
- Tabela 15 - Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com *cross-validation* de 10-folds, utilizando o vetor de características com os cinco atributos de maiores valores de A_z calculados para a análise “FVC x Normal”. Em negrito estão os valores das médias aritméticas dos resultados das classificações..... (68)
- Tabela 16 - Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com *cross-validation* de 10-folds, utilizando o vetor de características com os cinco atributos de maiores valores de A_z calculados para a análise “FVC Maligna x FVC Benigna”. Em negrito estão os valores das médias aritméticas dos resultados das classificações..... (69)
- Tabela 17 - Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com *cross-validation* de 10-folds, utilizando o vetor de características com os dez atributos de maiores valores de A_z calculados para a análise “FVC x Normal”. Em negrito estão os valores das médias aritméticas dos resultados das classificações..... (69)
- Tabela 18 - Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com *cross-validation* de 10-folds, utilizando o vetor de características com os dez atributos de maiores valores de A_z calculados para a análise “FVC Maligna x FVC Benigna”. Em negrito estão os valores das médias aritméticas dos resultados das classificações..... (70)
- Tabela 19 - Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com *cross-validation* de 10-folds, utilizando o vetor de características específico de cada classificador, obtido na seleção de atributo com o método *Wrapper*, calculados para a análise “FVC x Normal”. Em negrito estão os valores das

médias aritméticas dos resultados das classificações e os valores em (?) são valores não obtidos..... (70)

Tabela 20 - Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com *cross-validation* de 10-folds, utilizando o vetor de características específico de cada classificador, obtido na seleção de atributo com o método *Wrapper*, calculados para a análise “FVC Maligna x FVC Benigna”. Em negrito estão os valores das médias aritméticas dos resultados das classificações e os valores em (?) são valores não obtidos..... (71)

Tabela 21 - Acurácia e AUROC de cada classificação, com *cross-validation* de 10-folds, utilizando o vetor de características específico de cada classificador, obtido na seleção de atributo com o método *Wrapper*, calculados para a análise “FVC Maligna x FVC Benigna x Normal”. Em negrito estão os valores das médias aritméticas dos resultados das classificações e os valores em (?) são valores não obtidos..... (71)

Tabela 22 - Resultados de sensibilidade e especificidade, com seus respectivos intervalos de confiança de 95%, para as classificações feitas pelos radiologistas e pelos classificadores para as análises “FVC x Normal”..... (72)

Tabela 23 - Resultados de sensibilidade e especificidade, com seus respectivos intervalos de confiança de 95%, para as classificações feitas pelos radiologistas e pelos classificadores para as análises “FVC Maligna x FVC Benigna”..... (73)

LISTA DE GRÁFICOS

- Gráfico 1 - Valores de A_z para cada atributo na análise FVC x Normal. Em azul são os atributos de Níveis de Cinza, em vermelho os atributos de textura e em verde os atributos de forma. A linha em preto representa a linha exponencial de tendência..... (58)
- Gráfico 2 - Valores de A_z para cada atributo na análise FVC Maligna x FVC Benigna. Em azul são os atributos de Níveis de Cinza, em vermelho os atributos de textura e em verde os atributos de forma. A linha em preto representa a linha exponencial de tendência..... (60)
- Gráfico 3 - Ilustração das curvas ROC dos cinco atributos apresentados na Tabela 6 para a classificação de FVC x Normal..... (61)
- Gráfico 4 - Ilustração das curvas ROC dos cinco atributos apresentados na Tabela 7 para a classificação de FVC Maligna x FVC Benigna..... (62)
- Gráfico 5 - Ilustração dos resultados de sensibilidade, com intervalos de confiança de 95%, para “FVC x Normal” das avaliações dos três radiologistas e dos classificadores com método automático *Wrapper* de seleção de atributos de acordo com a Tabela 23..... (72)
- Gráfico 6 - Ilustração dos resultados de especificidade, com intervalos de confiança de 95%, para “FVC x Normal” das avaliações dos três radiologistas e dos classificadores com método automático *Wrapper* de seleção de atributos de acordo com a Tabela 23..... (73)
- Gráfico 7 - Ilustração dos resultados de sensibilidade, com intervalos de confiança de 95%, para “FVC Maligna x FVC Benigna” das avaliações dos três radiologistas e dos classificadores com método automático *Wrapper* de seleção de atributos de acordo com a Tabela 24..... (74)
- Gráfico 8 - Ilustração dos resultados de especificidade, com intervalos de confiança de 95%, para “FVC Maligna x FVC Benigna” das avaliações dos três radiologistas e dos

classificadores com método automático <i>Wrapper</i> de seleção de atributos de acordo com a Tabela 24.....	(74)
---	------

LISTA DE ABREVIATURAS

FVC	Fratura vertebral por compressão
IRM	Imagem por ressonância magnética
RM	Ressonância magnética
CAD	Sistema de diagnóstico auxiliado por computador
DXA	Absorciometria por duplo feixe de raio-X
TC	Tomografia computadorizada
ROI	Região de interesse, do inglês “ <i>region of interest</i> ”
RGB	Sistema de cores vermelho, verde, azul
CMYK	Sistema de cores ciano, magenta, amarelo, preto
v_p	Verdadeiro positivo
v_n	Verdadeiro negativo
f_p	Falso positivo
f_n	Falso negativo
acc	Acurácia
$sens$	Sensibilidade
TVP	Taxa de verdadeiros positivos
$spec$	Especificidade
ROC	Receiver Operating Characteristic
TFP	Taxa de falsos positivos
AUROC	Área sob a curva ROC
J48	Classificador de árvore de decisão
k -NN	Classificador <i>k-nearest neighbor</i> (<i>k</i> vizinhos mais próximos)
d_E	Distância Euclidiana

RBF	<i>Network Classificador Radial basis function network</i>
SVM	<i>Classificador Support vector machine</i>
CEP	Comitê de ética em pesquisa
HCFMRP-USP	Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo
TCLE	Termo de Consentimento Livre e Esclarecido
RIS	Sistema de informação de Radiologia
TE	Tempo de eco
TR	Tempo de repetição
FOV	Campo de visão, do inglês “ <i>field of view</i> ”
DICOM	<i>Digital Imaging and Communication in Medicine</i>
TIFF	<i>Tagged Image File Format</i>
HIS	Sistema de informação hospitalar
<i>cv</i>	Atributo de níveis de cinza: coeficiente de variação
<i>skew</i>	Atributo de níveis de cinza: <i>Skewness</i>
<i>kurt</i>	Atributo de níveis de cinza: <i>Kurtosis</i>
<i>MC</i>	Matriz de coocorrência dos níveis de cinza
h_1	Atributo de textura 1: segundo momento angular (energia)
h_2	Atributo de textura 2: contraste
h_3	Atributo de textura 3: correlação
h_4	Atributo de textura 4: soma de quadrados: variância
h_5	Atributo de textura 5: diferença inversa de momento
h_6	Atributo de textura 6: média da soma
h_7	Atributo de textura 7: variância da soma
h_8	Atributo de textura 8: entropia da soma

h_9	Atributo de textura 9: entropia
h_{10}	Atributo de textura 10: diferença da variância
h_{11}	Atributo de textura 11: diferença da entropia
h_{12}	Atributo de textura 12: medidas de informação de correlação 1
h_{13}	Atributo de textura 13: medidas de informação de correlação 2
h_{14}	Atributo de textura 14: coeficiente de máxima correlação
C_o	Atributo de forma: Compacidade
CD	Atributo de forma: deficiência convexa
FDF	Atributo de forma: descritor de Fourier
M_1	Atributo de forma: momento de Hu 1
M_2	Atributo de forma: momento de Hu 2
M_3	Atributo de forma: momento de Hu 3
M_4	Atributo de forma: momento de Hu 4
M_5	Atributo de forma: momento de Hu 5
M_6	Atributo de forma: momento de Hu 6
M_7	Atributo de forma: momento de Hu 7
μ_{pq}	Momentos centrais
m_{pq}	Momentos
mm	Milímetros
$meanDCB$	Atributo de forma: média da distancia do centroide vertebral a borda
$stdDCB$	Atributo de forma: desvio padrão da distancia do centroide vertebral a borda
$varDCB$	Atributo de forma: variância da distancia do centroide vertebral a borda
$cvDCB$	Atributo de forma: coeficiente de variação da distancia do centroide vertebral a borda
$skewDCB$	Atributo de forma: <i>skewness</i> da distância do centroide vertebral a borda

$kurtDCB$	Atributo de forma: <i>kurtosis</i> da distância do centroide vertebral a borda
mf	Atributo de forma: combinação dos momentos de quarta e segunda da distancia do centroide vertebral a borda
md	Média das distâncias Euclidianas
θ	Ângulo de inclinação do eixo principal
h_{si}	Altura vertebral na direção superior-inferior
w_{ap}	Largura vertebral na direção anterior-posterior
$meanLineV$	Atributo de forma: média de h_{si}
$stdLineV$	Atributo de forma: desvio padrão de h_{si}
$varLineV$	Atributo de forma: variância de h_{si}
$cvLineV$	Atributo de forma: coeficiente de variação de h_{si}
$skewLineV$	Atributo de forma: <i>skewness</i> de h_{si}
$kurtLineV$	Atributo de forma: <i>kurtosis</i> de h_{si}
$meanLineH$	Atributo de forma: média de w_{ap}
$stdLineH$	Atributo de forma: desvio padrão de w_{ap}
$varLineH$	Atributo de forma: variância de w_{ap}
$cvLineH$	Atributo de forma: coeficiente de variação de w_{ap}
$skewLineH$	Atributo de forma: <i>skewness</i> de w_{ap}
$kurtLineH$	Atributo de forma: <i>kurtosis</i> de w_{ap}
$meanLineHV$	Atributo de forma: media da união de h_{si} e w_{ap}
$stdLineHV$	Atributo de forma: desvio padrão da união de h_{si} e w_{ap}
$varLineHV$	Atributo de forma: variância da união de h_{si} e w_{ap}
$cvLineHV$	Atributo de forma: coeficiente de variação da união de h_{si} e w_{ap}
$skewLineHV$	Atributo de forma: <i>skewness</i> da união de h_{si} e w_{ap}

<i>kurtLineHV</i>	Atributo de forma: <i>kurtosis</i> da união de h_{si} e w_{ap}
MA	Média de h_{si} da porção anterior do corpo vertebral
MC	Média de h_{si} da porção central do corpo vertebral
MP	Média de h_{si} da porção posterior do corpo vertebral
<i>uAC</i>	Atributo de forma: MA-MC
<i>uCP</i>	Atributo de forma: MC-MP
<i>uPA</i>	Atributo de forma: MP-MA

RESUMO

CLASSIFICAÇÃO SEMIAUTOMÁTICA DE FRATURAS VERTEBRAIS BENIGNAS E MALIGNAS EM IMAGENS DE RESSONÂNCIA MAGNÉTICA

Propósito:

Fraturas vertebrais por compressão (FVCs) são caracterizadas por colapso parcial de corpos vertebrais. Elas tipicamente ocorrem na população idosa de forma não traumática ou por trauma de baixa energia, podendo ser secundárias a fragilidade causada pela osteoporose (FVCs benignas) ou metástases vertebrais (FVCs malignas). Nosso trabalho tem o objetivo de detectar a presença de FVCs e de classifica-las como FVC maligna ou FVC benigna utilizando técnicas de processamento de imagens e aprendizado de máquinas em imagens ponderadas em T1 obtidas em ressonância magnética (RM).

Materiais e Métodos:

Foram utilizados os planos sagitais medianos das RMs da coluna lombar de 63 pacientes (38 mulheres e 25 homens) previamente diagnosticados com FVCs. Os corpos vertebrais lombares foram segmentados manualmente. Atributos de análise de níveis de cinza foram calculados do histograma dos corpos vertebrais. Foram extraídos também atributos de textura para analisar a distribuição dos níveis de cinza e atributos de forma para analisar o formato dos corpos vertebrais. No total, 102 FVCs lombares (53 benignas e 49 malignas) e 89 corpos vertebrais lombares foram analisados. Após a aplicação de métodos de seleção de atributos nos vetores de características, foram realizadas classificações com os classificadores *k-nearest-neighbor* (*k*-NN), uma rede neural artificial com função de base radial (*RBF network*), *naïve* Bayes, J48 e *Support Vector Machine* (SVM). O padrão de referência para calcular o desempenho diagnóstico do sistema desenvolvido foi uma classificação obtida do prontuário médico eletrônico com o diagnóstico final de cada caso, incluindo no mínimo informações a respeito de

biópsia para as FVC lesões malignas e acompanhamento clínico e laboratorial para as FVCs benignas. Três radiologistas classificaram os mesmos casos analisando as mesmas regiões de interesse (ROIs) que os classificadores e uma comparação entre classificadores e radiologistas foi realizada.

Resultados:

Os resultados obtidos pelos classificadores mostraram uma área abaixo da curva *receiver operating characteristic (AUROC)* de 0,984 para distinguir entre corpos vertebrais com FVC e normais e *AUROC* de 0,930 para discriminar entre FVC benigna e FVC maligna.

Conclusão:

Nosso método alcançou ótimos resultados na classificação de corpos vertebrais sem fratura, corpos vertebrais com fratura por osteoporose e corpos vertebrais com fraturas secundárias a doença metastática. Nossos resultados foram estatisticamente equivalentes ao de médicos radiologistas e se mostraram promissores na assistência em diagnóstico de FVCs.

Palavras chave: Diagnóstico auxiliado por computador, Fratura vertebral por compressão, Imagens de ressonância magnética, Processamento de imagens.

ABSTRACT

SEMI-AUTOMATIC CLASSIFICATION OF BENIGN AND MALIGNANT VERTEBRAL FRACTURES IN MAGNETIC RESONANCE IMAGING

Purpose:

Vertebral compression fractures (VCFs) result in partial collapse of vertebral bodies. They usually are nontraumatic or occur with low-energy trauma in the elderly secondary to different etiologies, such as insufficiency fractures of bone fragility in osteoporosis (benign fractures) or vertebral metastasis (malignant fractures). Our study aims to detect the presence of VCFs and classify them as malignant and benign using image processing techniques and machine learning classifiers in T1-weighted magnetic resonance images (MRI).

Materials and methods:

*We used the median sagittal planes of lumbar spine MRIs from 63 patients (38 women and 25 men) previously diagnosed with VCFs. The lumbar vertebral bodies were manually segmented and statistical features of gray levels were computed from the histogram. We also extracted texture features to analyze the gray-level distribution, and shape features to analyze the contours of the vertebral bodies. In total, 102 lumbar VCFs (53 benign and 49 malignant) and 89 normal lumbar vertebral bodies were analyzed. After run feature selection methods to the vector of features, the *k*-nearest-neighbor (*k*-NN), neural network with radial basis functions (RBF network), a naïve Bayes classifier, J48, and Support Vector Machine (SVM) were used for classification. We compared the classification obtained by these classifiers with the final diagnosis of each case, including biopsy for the malignant fractures and clinical and laboratory follow up for the benign fractures. Furthermore, three voluntary radiologists classified the same cases analyzing the same regions of interests (ROIs) used by the classifiers and a comparison between the classifiers and the radiologists was done.*

Results:

The results obtained by the classifiers showed an area under the receiver operating characteristic curve (AUROC) of 0.984 in distinguishing between normal and fractured vertebral bodies, and AUROC of 0.930 in discriminating between benign and malignant VCFs.

Conclusion:

Our method reached great results in the classification of vertebral bodies without fractures, vertebral bodies with fractures due to osteoporosis and vertebral bodies with fractures due to metastatic diseases. Our results were statistically equivalent to the results of the classifications made by radiologists and they showed to be promising in diagnosis assisting of VCFs.

Keywords: computer aided diagnosis, image processing, magnetic resonance imaging, vertebral compression fractures.

SUMÁRIO

1. INTRODUÇÃO.....	1
1.1. Fraturas vertebrais por compressão	1
1.2. Diagnóstico auxiliado por Computador.....	4
1.3. Revisão bibliográfica de estudos em FVCs	5
1.4. Processamento de imagens médicas	8
1.5. Aprendizado de máquinas	12
2. JUSTIFICATIVA.....	19
3. HIPÓTESE.....	20
4. OBJETIVO	21
5. MATERIAIS E MÉTODOS	22
5.1. Aprovação do Comitê de Ética em Pesquisa	22
5.2. Base de dados.....	22
5.4. Desenvolvimento do software	25
5.5. Classificação e estatística.....	48
6. RESULTADOS.....	56
6.1. Estatística dos atributos: p-valores e A_z	56
6.2. Seleções de atributos	61
6.3. Classificação.....	67
7. DISCUSSÃO.....	75
7.1. Seleção de atributos	75
7.2. Classificação dos algoritmos.....	79
7.3. Classificação dos radiologistas e computador.....	84
8. CONCLUSÃO	89
9. REFERÊNCIAS	90
Anexo A	95
Anexo B	96
Anexo C	97
Anexo D	98

Anexo E	99
Anexo F	100
Anexo G.....	101

1. INTRODUÇÃO

1.1. Fraturas vertebrais por compressão

A fratura vertebral por compressão (FVC) se manifesta nos exames de imagem como colapso parcial do corpo vertebral. Na rotina clínica, o médico radiologista em geral não terá dúvidas sobre a etiologia de uma FVC traumática, pois, o paciente apresentará um histórico de trauma recente. Porém, quando um paciente desenvolve um colapso vertebral recente sem histórico de trauma, ou quando houve trauma de baixa energia, como uma queda da cama, sua etiologia precisa ser investigada e, conseqüentemente, dificuldades no diagnóstico podem ocorrer. Esse potencial dilema no diagnóstico é comum especialmente na população idosa devido à alta incidência e alta prevalência de FVCs secundárias a fragilidade óssea quando comparado com a população jovem. A doença metabólica óssea mais comum é a osteoporose e as FVCs são as fraturas mais comuns dentre as fraturas causadas pela osteoporose. Com o aumento da expectativa de vida as FVCs secundárias à osteoporose tendem a ser cada vez mais comuns e estas fraturas tem sido motivo de estudos científicos (KONDO, 2008; OEI *et al.*, 2013). Por outro lado, a população idosa possui também uma alta incidência e prevalência de FVCs relacionadas ao câncer metastático afetando ossos (TEHRANZADEH e TAO, 2004). Clinicamente, uma FVC secundária à osteoporose é dita como FVC benigna e a FVC causada por metástase óssea é considerada como FVC maligna (TEHRANZADEH e TAO, 2004). As queixas clínicas de pacientes referentes a ambos os tipos de FVC, benigna e maligna, podem ser similares, porém a diferenciação entre ambas é fundamental para o seu correto diagnóstico e tratamento.

Atualmente, há reconhecida necessidade de redução da dosagem de radiação ionizante a qual os pacientes em procedimentos de imagens médicas são expostos (SEMELKA *et al.*, 2007). Além da vantagem de não utilizar radiação ionizante, a modalidade de imagem por ressonância magnética (IRM) é a mais confiável e amplamente utilizada para alterações da coluna vertebral. IRM é efetiva na detecção antecipada de fraturas vertebrais (UETANI *et al.*,

2004; PRASAD e SCHIFF, 2005). A distribuição da intensidade de sinal no corpo vertebral é um importante critério para discriminar FVCs benignas e malignas em IRM, porém a análise do formato do corpo vertebral pode também ajudar nessa discriminação (CUENOD *et al.*, 1996).

Um colapso vertebral maligno tipicamente exibe baixa intensidade de sinal em todo o corpo vertebral em IRM nas imagens ponderadas em T1. FVCs por osteoporose, em contrapartida, caracteristicamente demonstram preservação parcial do sinal normal de gordura da medula óssea do corpo vertebral nas IRM ponderadas em T1 (CUENOD *et al.*, 1996). Na rotina clínica, para avaliar a possível presença de doenças malignas infiltrativas na medula óssea vertebral, dentre outros parâmetros, o radiologista compara a intensidade de sinal do corpo vertebral com a intensidade de sinal dos discos intervertebrais. Geralmente, corpos vertebrais normais exibem uma intensidade de sinal maior que a intensidade de sinal dos discos intervertebrais em IRM com ponderação em T1, devido à adiposidade da medula óssea. A medula óssea do corpo vertebral com intensidade de sinal similar ao de um disco intervertebral vizinho favorece a suspeita de malignidade (KAPLAN *et al.*, 2001).

Link *et al.* (2005) descreveu sinais comuns de FVCs benignas e malignas em IRM. FVCs benignas podem apresentar cortex posterior côncavo, retropulsão de fragmentos posteriores no canal espinhal, sinal medular preservado em IRM com ponderação em T1 e vertebra isointensa em IRM com ponderação T1 e T2 com contraste de gadolínio, além de baixa intensidade de sinal nos pedículos na ponderação T1.

Vértebras não fraturadas, geralmente, apresentam forma aproximadamente retangular nas imagens do plano sagital. Porém, frequentemente vértebras de pacientes idosos apresentam pequenas proliferações ósseas marginais conhecidas como osteófitos que podem causar pequenas alterações da morfologia. Tais modificações de forma podem potencialmente confundir análises com fatores típicos de forma como compacidade e deficiência convexa (RANGAYYAN, 2005; CABRAL e RANGAYYAN, 2012; RAGUSO *et al.*, 2010).

FVCs causam diferentes alterações no formato do corpo vertebral. FVCs malignas podem resultar em um abaulamento da parede posterior do corpo vertebral, com contornos mais suaves, devido ao abaulamento do tecido neoplásico. Isso pode ocorrer em conjunto com uma deformação côncava dos platôs vertebrais. No caso de FVCs benignas, os platôs vertebrais podem adquirir um formato côncavo mais acentuado e a impactação do osso subcondral pode resultar em contornos ásperos com reentrâncias. FVCs benignas podem causar retropulsão de fragmentos da parede posterior com contornos angulados ou irregulares (CUENOD *et al.*, 1996; JUNG *et al.*, 2003).

Na prática clínica, há casos de FVCs malignas e benignas que não seguem as características típicas citadas, o que exige outros métodos de investigação precisos e reprodutíveis para FVCs como análise de histórico do paciente, exames de biópsia e o auxílio de técnicas computacionais. A Figura 1 traz três exames de ressonância magnética (RM).

Figura 1 – Ilustração de três IRM da coluna vertebral de diferentes pacientes em cortes sagitais medianos. Em (a) os corpos vertebrais estão íntegros não havendo presença de fraturas vertebrais. (b) apresenta duas FVCs nos corpos vertebrais L1 e L4 de características benignas. (c) apresenta duas FVC nos corpos vertebrais L1 e L5 secundárias a metástase óssea.

1.2. Diagnóstico auxiliado por Computador

Sabendo que médicos radiologistas e outros profissionais de saúde são altamente treinados e qualificados em análise e diagnóstico por imagens médicas, por que esses profissionais precisam do computador para a análise dessas imagens?

Seres humanos são altamente habilidosos e rápidos na análise de padrões visuais, mas são geralmente lentos em operações aritméticas com grande quantidade de valores. Por exemplo, uma IRM de coluna vertebral com resolução de 960×960 -pixel contém $92,16 \times 10^4$ pixels. Se essa imagem necessitar de um processamento para a remoção de ruídos ou extração de parâmetros, não será prático para um ser humano realizar essa tarefa. Um computador é capaz de realizar milhares de operações matemáticas em segundos, tornando a tarefa viável. Além disso, os seres humanos podem ser afetados pela fadiga, fatores do ambiente, estresse e por isso ficam mais suscetíveis a cometerem erros. Dessa forma, em situações limítrofes de cansaço alterações presentes nas imagens podem passar despercebidos pelos radiologistas. Os computadores, no entanto, são designados a realizarem tarefas específicas e repetitivas.

Análises feitas por seres humanos são comumente subjetivas e qualitativas. Ao fazer uma análise comparativa de imagens, o profissional de saúde fornecerá tipicamente uma resposta qualitativa ou semiquantitativa. Além disso, essa análise está sujeita a variação intraobservador, devido a um possível descuido ou falta do conhecimento específico, e interobservador, devido a variações de treinamento e nível de conhecimento ou competência, sendo ambas as variações passíveis de variar com o tempo.

A derivação de características numéricas ou quantitativas de imagens certamente demandará o uso de computadores. A lógica de diagnóstico médico via análise de imagens pode ser objetivamente codificada e consistentemente aplicada na rotina de tarefas repetitivas. Porém, deve ser enfatizado que o objetivo final da análise de imagem biomédica deve ser o diagnóstico auxiliado por computador e não o diagnóstico automatizado. Os resultados de

análise de imagens devem ser integrados com outros sinais clínicos, sintomas e informações pelo médico especialista. Provavelmente a intuição do médico especialista tem importante papel para chegar do diagnóstico final correto.

Portanto, a expectativa é de que análises quantitativas e objetivas de imagens médicas facilitadas por aplicações e softwares de computador possam conduzir o especialista a uma decisão diagnóstica mais acurada. Softwares como esse são conhecidos como sistemas de diagnóstico auxiliado por computador (CAD, em inglês: “*Computer-aided Diagnosis system*”) (RANGAYYAN, 2005).

1.3. Revisão bibliográfica de estudos em FVCs

Genant et al. (1993) propuseram um método de classificação para avaliação semiquantitativa de fraturas vertebrais osteoporóticas. Esse método compara as alturas da porção anterior, central e posterior dos corpos vertebrais. Corpos vertebrais normais são dados como grau zero. Deformidades discretas, com de redução da altura entre 20% e 25% em qualquer porção do corpo vertebral, são classificadas como grau I. O grau II é atribuído às deformidades moderadas com redução de altura vertebral em qualquer porção de 26% a 40%. Reduções de altura vertebral, em qualquer porção, acima de 40% são classificadas como deformidades graves e classificadas como grau III. A Figura 2 foi retirada e traduzida de Genant et al. (1993) a qual ilustra os graus de deformidade dos corpos vertebrais.

Grados et al. (2009) apresentaram uma revisão de 149 artigos com descrição e avaliação de fraturas vertebrais osteoporóticas com radiografias padrões ou absorciometria por duplo feixe de raio-X (DXA, em inglês: “*dual-energy X-ray absorptiometry*”) e o papel de cada método na prática clínica. O método de classificação de Genant et al. (1993) para avaliação de FVCs utilizando radiografias padrões fornece informação na gravidade e prognóstico da osteoporose, enquanto DXA para a mensuração da densidade mineral óssea pode detectar fracturas vertebrais mesmo em pacientes assintomáticos.

Figura 2 – Deformidades dos corpos vertebrais em graus de lesões de acordo com Genant et al. (1993).

Guglielmi et al. (2009) conduziram estudo morfométrico em FVCs benignas em radiografias e avaliaram uma ferramenta de um software que analisa a imagem semiautomaticamente utilizando um protocolo de morfometria de 6 pontos e um processo manual de acordo com as medidas de forma e tamanho de Genant et al. (1993). O tempo consumido pelo processo semiautomático foi significativamente menor que o tempo consumido pelo processo manual. A abordagem semiautomática resultou em um erro médio e desvio padrão de $2,50 \pm 0,72\%$ e $2,16 \pm 0,5\%$ para medições anterior-posterior e superior-inferior, respectivamente. Eles também sugerem que a deformidade vertebral nem sempre indica uma fratura do corpo vertebral, porém uma fratura vertebral sempre está acompanhada de uma deformidade vertebral e também que a combinação de métodos visuais semiquantitativos e métodos morfométricos quantitativos podem ser a melhor abordagem para o diagnóstico de fraturas vertebrais.

Ribeiro et al. (2012) propuseram procedimentos para detecção de FVCs em radiografias laterais da coluna lombar utilizando filtros de Gabor e uma rede neural artificial para extração dos platôs superior e inferior de cada corpo vertebral e aplicou o método de Gentant et al.

(1993) para classificação utilizando medidas de altura dos corpos vertebrais. Seus resultados indicaram sensibilidade de até 78% e especificidade de 95%.

Kasai et al. (2006) desenvolveram um sistema automático de CAD para fraturas vertebrais baseado em imagens de raio-x do tórax. Comparando com análises manuais feitas por radiologistas, eles obtiveram acurácia variando entre 70,9% e 76,6% na detecção de fraturas vertebrais.

Al-Helo et al. (2013) desenvolveram um sistema CAD para detecção de fraturas em corpos vertebrais em imagens de tomografia computadorizada (TC). Utilizando modelos ativos de forma e fluxo ativo de vetores gradientes de contornos, eles obtiveram 99% de sensibilidade e 87,5% de especificidade com o método *k*-means. Ghosh et al. (2011) propuseram um sistema automático de CAD completo para detecção de FVCs da coluna lombar em imagens de TC utilizando três medidas de variação de altura de corpos vertebrais como características. Eles obtiveram uma acurácia de 97,33%. Enquanto TC é a modalidade preferida para pacientes com histórico de trauma, IRM é utilizada na maioria dos outros cenários clínicos de investigação de doenças na coluna vertebral (GHOSH et al., 2011).

Os trabalhos anteriormente revistos acima tem em comum o uso de modalidades de imagem com radiação ionizante. Arevalo-Perez et al. (2015) estudaram o uso de contraste dinâmico de realce em IRM a fim de auxiliar na distinção não invasiva entre fraturas vertebrais benignas e patológicas. Eles revisaram casos de pacientes com fraturas vertebrais os quais foram submetidos à IRM com contraste dinâmico de realce e biópsia. Os resultados indicaram que fraturas patológicas obtiveram parâmetros de perfusão significativamente maiores ($p < 0,01$) comparados com fraturas benignas. Khan et al. (2015) estudaram a variação no envelhecimento da coluna lombar usando análise fatorial e principais componentes em IRM. Eles usaram o MATLAB Statistical Toolbox (MATLAB, 2014) para computar os principais componentes utilizando a técnica de variância, e encontraram que as intensidades de sinais de

discos tem forte correlação com envelhecimento natural e que as alturas dos discos e corpos vertebrais são correlacionadas.

1.4. Processamento de imagens médicas

Como visto anteriormente, sistemas CAD têm exercido papéis importantes como ferramenta auxiliar no diagnóstico médico como visualização, interpretação e quantificação de imagens médicas. Para isto diversas técnicas de processamento de imagens podem ser aplicadas, tais como: segmentação, máscara e extração de características da imagem.

1.4.1. Segmentação

A segmentação consiste basicamente da separação de estruturas ou regiões de interesse (ROIs, em inglês: “*regions of interest*”) da imagem através de *labels* que delimitam as mesmas. É uma das técnicas de processamento de imagens mais utilizadas e possui aplicabilidade nas mais diversas subáreas da medicina como delineação de ROIs, como um órgão, tumor ou estrutura como um corpo vertebral, a fim analisá-las individualmente.

A segmentação pode ser realizada em *softwares* de forma manual, semiautomática ou automática. As duas últimas formas têm despertado o interesse de instituições de pesquisa (TANG e PAULI, 2011; DUARTE BARBIERI *et al.*, 2015).

1.4.2. Máscara

A máscara é uma imagem gerada por segmentação de ROIs ou por um algoritmo de *threshold*, geralmente binária, a qual é aplicada a uma determinada imagem a fim de se considerar apenas suas ROIs. Com ela é possível extrair em IRM apenas os corpos vertebrais lombares, por exemplo.

1.4.3. Extração de características

Em sistemas computacionais, uma imagem digital é representada por uma matriz. Cada célula ou ponto dessa matriz é chamado de “pixel” e seu valor determina a intensidade da cor ou nível de cinza nesse ponto. Dessa forma, é possível se utilizar funções matemáticas e estatísticas, chamadas de descritores, com o objetivo de extrair características da imagem. O conjunto de descritores ou atributos calculados é chamado assinatura, ou vetor de características.

O vetor de características da imagem ou ROI possibilita a realização de classificação que satisfaça um determinado critério de similaridade. As características da imagem tem relação com aspectos de forma, textura e cor ou nível de cinza.

i. Forma

As características de forma se referem ao formato geométrico de uma ROI ou objeto presente na imagem. Ao analisar novamente a Figura 2, nota-se que o corpo vertebral saudável apresenta um formato próximo a um retângulo, enquanto os corpos vertebrais com fraturas apresentam alterações importantes em sua forma. Descritores de forma necessitam ser invariantes à rotação, translação e escala para serem comparáveis.

ii. Níveis de Cinza e cor

A cor é uma propriedade da percepção visual derivada dos diferentes espectros e intensidades da luz. Na representação computacional são utilizados sistemas de cores, como o RGB (em inglês: “*red* (vermelho), *green* (verde), *blue* (azul)”) e o CMYK (em inglês: “*cian* (ciano), *magenta* (magenta), *yellow* (amarelo), *key: black* (preto)”), que combinam características primárias para formar as diferentes cores. Comumente, em IRM, após serem adquiridas pelo aparelho, as imagens são geradas computacionalmente em níveis de cinza. Uma forma de representação da imagem quanto aos níveis de cinza é através do seu

histograma. O histograma fornece a quantidade de pixels na imagem correspondente a cada valor de nível de cinza dentro de uma escala. Por exemplo, em uma imagem de 8-bits/pixel, a escala de nível de cinza é de $[0, 255]$ sendo o valor zero correspondente a mais baixa intensidade de luz, ou seja, a cor preta; o valor 255 correspondente a mais alta intensidade de luz, ou seja, a cor branca; e entre os limites estão às variações de cinza. Com o histograma, é possível calcular descritores estatísticos dos níveis de cinza da imagem. A Figura 3 traz o histograma de níveis de cinza de uma de ROI de um corpo vertebral.

Figura 3 – Histograma de níveis de cinza de uma ROI de 8-bits/pixel de um corpo vertebral com FVC benigna. Note que apenas a região do corpo vertebral foi utilizada para o cálculo do histograma, excluindo o fundo preto.

iii. Textura

Embora não haja uma definição unânime, em imagens, a textura pode ser entendida como a distribuição espacial dos pixels de uma região formando um padrão visual. Três abordagens são comumente aplicadas para caracterização da textura: a estrutural, que considera arranjos de pixels como estrutura primitiva de textura; a espectral, que por meio de transformadas, busca fazer a representação das imagens no domínio de frequência; a

estatística, que considera a distribuição dos valores dos pixels e a relação entre eles, e esse inter-relacionamento dos valores de cada pixel com seus vizinhos são representados pela matriz de coocorrência dos níveis de cinza calculada pela relação de dois parâmetros: distância e o ângulo entre os níveis de cinza. A fim de ilustrar a matriz de coocorrência, a Figura 4 traz uma imagem qualquer de 5×5 pixels na escala de três níveis de cinza [0, 2], sendo os pixels de valor = 0 representado pelo preto, valor = 1 representado pelo cinza e o valor = 2 representado pelo branco. Como a imagem está na escala de três níveis de cinza, sua matriz de coocorrência será uma matriz 3×3 pixels. Uma imagem de 8-*bits*/pixel com 256 níveis de cinza teria uma matriz de coocorrência de 256×256 pixels. A matriz de coocorrência da imagem apresentada na Figura 4 com distância = 1 dos pixels e um ângulo de $\pm 45^\circ$ está ilustrada na Figura 5.

Figura 4 – Imagem 5×5 em uma escala com três níveis de cinza: (a) representação visual; (b) representação computacional.

Figura 5 – Em destaque está a matriz de coocorrência de níveis de cinza da imagem da Figura 4 com distância = 1 e ângulo = $\pm 45^\circ$.

Em imagens médicas, é possível calcular a matriz de coocorrência de níveis de cinza de uma ROI de um corpo vertebral e, com ela, extrair descritores de textura como os de Haralick et al. (1973).

1.5. Aprendizado de máquinas

O propósito da análise de imagens biomédicas é classificar suas características que foram detectadas e extraídas para uma ou mais categorias. O reconhecimento de padrões ou classificação pode ser definido como a categorização dos dados de entrada para classes identificáveis via extração de características ou atributos caracterizadores. Em análise de imagens médicas, depois da extração quantitativa das características de uma imagem ou ROI, esta pode ser representada por um vetor de características. Quando cada valor das características pertence aos números reais, esse vetor de características será um ponto em um espaço Euclidiano n -dimensional. Dessa forma, a expectativa é de que os vetores de características de objetos similares tenham distâncias semelhantes nesse espaço Euclidiano (RANGAYAN, 2005).

Considere uma situação a qual representa o uso de um classificador, ou seja, um algoritmo de aprendizado de máquina para a classificação de presença ou ausência de uma

FVC. Casos de FVCs são classificados como positivos e casos normais são classificados como negativos. O classificador deve ser treinado e testado para essa classificação. O treinamento é a etapa de aprendizagem do classificador, onde este recebe uma amostra de vetores características de casos positivos e negativos. O teste é a etapa onde o classificador receberá uma nova amostra de vetores características de casos positivos e negativos a fim de classificá-los de acordo com o aprendizado recebido na etapa de treinamento.

Em casos em que a amostra não é suficientemente grande é comum o emprego de técnicas de amostragem. *Cross-validation* estratificada de *r-fold* (*stratified r-fold cross-validation*), por exemplo, é uma técnica de amostragem que divide aleatoriamente o conjunto em *r-folds* independentes, isto é, sem intersecções, de mesmo tamanho e mantendo a proporção entre o número de exemplos de cada classe (estratificação). O treinamento é realizado em $r - 1$ *folds* utilizando o remanescente para o teste. Este processo é realizado r vezes utilizando em cada interação um *fold* diferente como conjunto de teste.

As respostas do classificador devem ser validadas e essa validação possibilitará uma análise de seu desempenho para essa classificação a que foi treinado. Uma resposta verdadeira positiva (v_p) significa que o teste classificou corretamente um caso de FVC. Uma resposta verdadeira negativa (v_n) significa que o teste classificou corretamente um caso normal. Já respostas falso positivas (f_p), ou erro tipo I, significam que o teste classificou incorretamente casos normais, ou seja, o teste classificou determinado caso como FVC (positivo) quando a resposta correta seria normal (negativo). E respostas falso negativas (f_n), ou erro tipo II, significam que o teste classificou incorretamente casos de FVC, ou seja, o teste classificou determinado caso como normal (negativo) quando a resposta correta seria FVC (positivo).

Uma forma de representar esses resultados é através da matriz de confusão representada na Tabela 1. A diagonal principal da matriz de confusão representa o total de acerto do teste e $n = v_p + v_n + f_p + f_n$ representa o total de casos da amostra.

Tabela 24 – Exemplo de uma matriz de confusão

Resposta correta	Resposta do teste	
	FVC	Normal
FVC	v_p	f_n
Normal	f_p	v_n

Da matriz de confusão são extraídas três medidas importantes para a avaliação de classificadores: acurácia (acc), sensibilidade ($sens$), também chamada de taxa de verdadeiros positivos (TVP) ou *Recall*, e especificidade ($spec$):

$$acc = \frac{v_p + v_n}{n} \quad \dots\dots\dots (1)$$

$$TVP = Recall = sens = \frac{v_p}{v_p + f_n} \quad \dots\dots\dots (2)$$

$$spec = \frac{v_n}{v_n + f_p} \quad \dots\dots\dots (3)$$

Com a sensibilidade e a especificidade, pode-se obter a curva *Receiver Operating Characteristic* (ROC). O gráfico de uma curva ROC representa em seu eixo horizontal a taxa de f_p (TFP), onde $TFP = 1 - spec$, e em seu eixo vertical a sensibilidade ou TVP . Cada ponto do gráfico é obtido a partir de uma matriz de confusão determinada pelo *threshold* ou limiar de separação nas distribuições das classes. O limiar mínimo, portanto, será o ponto (0,0) da curva ROC, enquanto o limiar máximo será o ponto (1,1) da curva ROC.

Uma medida de desempenho do classificador retirada deste gráfico é a área sob a curva ROC (AUROC ou A_z , em inglês: “*area under ROC curve*”). O classificador ideal, para um

determinado conjunto de exemplos, seria aquele que possuísse $TFP = 0$ e $TVP = 1$, portanto $A_z = 1$.

Um software de aprendizado de máquinas frequentemente utilizado na literatura é o software WEKA (HALL *et al.*, 2009). Esse software possui pacotes com diversos classificadores, além de métodos de seleção de atributos ou características.

A presença de atributos irrelevantes no vetor de características geralmente confunde o algoritmo de aprendizagem e, conseqüentemente, a classificação pode vir a ter um desempenho inferior comparada à classificação de um algoritmo que recebe um vetor de características com o mínimo de atributos irrelevantes. O ideal seria montar o vetor de características com atributos baseados em estudos e discussões sobre o problema a ser resolvido. Mesmo assim alguns atributos podem não ser viáveis em algumas das análises de classificação. A fim de reduzir a quantidade desses atributos irrelevantes, métodos de seleções de atributos podem ser aplicados aos vetores de características antes das classificações.

Algoritmos de seleção de atributos selecionam automaticamente os atributos mais relevantes de um dado vetor de características seguindo uma metodologia. O método *Wrapper*, presente no software WEKA (HALL *et al.*, 2009), é uma abordagem com o objetivo de gerar um subconjunto otimizado do vetor de característica para um algoritmo de classificação específico. A abordagem *Wrapper* gera um subconjunto de atributos do vetor de características como candidato, executa o classificador específico com esse subconjunto como treinamento e o avalia com a ACC obtida. A busca pelos subconjuntos é feita utilizando uma função de busca e ela é feita até que o critério de parada ocorra (KOHAVI e JOHN, 1997).

Alguns classificadores clássicos na literatura estão presentes em pacotes do software WEKA:

i. Árvore de decisão J48

O algoritmo J48 é um algoritmo de árvore de decisão popular implementado na linguagem JAVA do algoritmo de árvore de decisão C4.5 (QUILAN, 1993). Ele é baseado na estratégia de *top-down* (atributos são organizados em hierarquias) e na estratégia recursiva dividir para conquistar. O algoritmo escolhe qual atributo será o *root* o qual será o primeiro atributo nó da hierarquia da árvore de decisão. A escolha é feita através da teoria da informação que é um cálculo de ganho (entropia/pureza) medido em *bits* cujo objetivo é selecionar um atributo que separe as classes de uma maneira mais otimizada e “pura” e formar a menor árvore possível. Esse nó é dividido em n ramos de intervalos de valores possíveis de *root*. O cálculo de ganho é feito recursivamente para a criação dos demais nós da árvore de decisão (QUILAN, 1993).

ii. *k-nearest neighbor (k-NN)*

O algoritmo *k-NN* utiliza a estratégia de similaridade dos objetos, por exemplo corpos vertebrais, através da distância Euclidiana (d_E) entre os k vizinhos mais próximos presentes no espaço Euclidiano. Ou seja, dado um vetor de características de um corpo vertebral a ser classificado por um classificador *k-NN*, sua classe será dada pela classe majoritária dos k vetores de características dos corpos vertebrais mais próximos presentes no espaço Euclidiano (AHA *et al.*, 1991).

$$d_E = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad \text{..... (4)}$$

Onde p e q são dois vetores de características e n o tamanho dos vetores, ou seja, a quantidade de atributos.

iii. Naïve Bayes

O classificador *naïve Bayes* vem da família dos classificadores de probabilidades baseados na aplicação do teorema de Bayes para determinação da probabilidade de um vetor de característica pertencer à determinada classe dado valores de seus atributos (HALL *et al.*, 2009; JOHN e LANGLEY, 1995). Esse classificador se provê efetivo em aplicações médicas (RISH, 2001).

$$P(Classe|ValorAtributo) = \frac{P(ValorAtributo|Classe)P(Classe)}{P(ValorAtributo)} \quad \dots\dots\dots (5)$$

iv. Radial basis function network (RBF network)

A RBF *network* é uma rede neural artificial que utiliza funções de base radial como funções de ativação (LEE *et al.*, 1999). Essa rede neural possui três camadas: camada de entrada (*input*), uma camada oculta intermediária (*hidden*) com uma função de base radial de ativação não linear e uma camada de saída linear (*output*). No caso da WEKA, a RBF *network* implementa uma rede neural de base radial Gaussiana normalizada. Ele utiliza o algoritmo de clusterização *k-means* para fornecer as funções de base e aprende uma regressão linear em cima disso. As Gaussianas simétricas multivariadas são ajustadas para os dados de cada cluster. A seguir está a função Gaussiana normalizada:

$$s(x) = \frac{\sum_{i=1}^N \lambda_i \phi(\|x - c_i\|)}{\sum_{i=1}^N \phi(\|x - c_i\|)} \quad \dots\dots\dots (6)$$

Onde $\phi(\|x - c_i\|) = \exp[-\gamma\|x - c_i\|^2]$ é a função de base radial, N é o número de neurônios na camada *hidden*, x é o vetor de entrada da rede neural (*input*), c_i é o vetor central do neurônio i , λ_i é o peso do neurônio i no neurônio linear de saída (*output*).

v. Support vector machine (SVM)

O SVM é essencialmente um classificador de duas classes que maximiza a largura da margem entre as classes, isto é, a área vazia em torno do hiperplano de separação definido da distância para as amostras de formação mais próximas. Pode ser alargado a problemas de múltiplas classes, resolvendo essencialmente vários problemas de duas classes (BURGES, 1998).

Dado um conjunto de treinamento que consiste em N pares de $(x_1, y_1), (x_2, y_2) \dots (x_n, y_n)$, onde x_i denotam as observações de formação n-dimensional e $y_i \in \{-1, 1\}$ são os rótulos de classificação correspondentes. O método SVM procura encontrar o hiperplano definido por

$$f(x) = (x \cdot w) + b = 0 \quad \dots\dots\dots (7)$$

que separa observações positivas e negativas com a margem máxima. O vetor w e o valor limite escalar b determinam a orientação e a posição do hiperplano de separação. Pode ser mostrado que o vetor solução W_{svm} é definido em termos de uma combinação linear das observações de formação, isto é,

$$W_{svm} = \sum_{i=1}^g \alpha_i y_i x_i \quad \dots\dots\dots (8)$$

onde α_i são coeficientes de Lagrange não negativos obtidos através da resolução de um problema de otimização quadrática com restrições de desigualdades lineares. Essas observações de treinamento x_i com α_i diferente de zero encontram-se no limite da margem e são chamados vetores de suporte.

Caso os padrões não sejam linearmente separáveis, uma das vantagens do SVM é que possibilita uma mudança da dimensionalidade através da aplicação de *kernels* não lineares.

2. JUSTIFICATIVA

Baseado nas observações descritas na sessão anterior fica evidente a importância de sistemas CAD na medicina e não identificamos trabalhos científicos na literatura sobre a classificação de FVCs benignas e malignas utilizando métodos de processamento de imagens para extração de características diagnósticas em imagens de RM.

3. HIPÓTESE

A hipótese do nosso trabalho foi que um método computadorizado conseguiria detectar FVCs em corpos vertebrais e classificar essas fraturas como malignas ou benignas utilizando técnicas de processamento de imagens e aprendizado de máquinas.

4. OBJETIVO

O objetivo principal do trabalho é:

- Desenvolver técnicas de processamento de imagens em RM para extrair características de corpos vertebrais da coluna lombar, detectar computacionalmente a presença de FVC e classifica-las como malignas ou benignas, utilizando algoritmos de aprendizado de máquina.

Os objetivos secundários são:

- Avaliar quais atributos e métodos de seleção de atributos proporcionam melhor desempenho diagnóstico para o algoritmo computacional.
- Comparar o desempenho do método auxiliado por computador com o desempenho de médicos radiologistas.

5. MATERIAIS E MÉTODOS

5.1. Aprovação do Comitê de Ética em Pesquisa

Esse trabalho obteve aprovação do Comitê de Ética em Pesquisa (CEP) do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto da Universidade de São Paulo (HCFMRP-USP), bem como a solicitação de dispensa do Termo de Consentimento Livre e Esclarecido (TCLE), de acordo com o processo de número 13414/2014. O documento de aprovação encontra-se no Anexo A.

5.2. Base de dados

Foi realizada uma busca no Sistema de Informação de Radiologia (RIS, em inglês: “*Radiology Information System*”) do hospital de nossa instituição, HCFMRP-USP, de pacientes em tratamento de fraturas vertebrais que realizaram exames de RM. Foram recuperados 220 prontuários de potenciais pacientes para o estudo de setembro de 2010 a março de 2014 (43 meses), e que foram examinados de forma consecutiva no setor de Radiologia de nossa instituição.

Foram excluídos casos de fraturas vertebrais secundárias a trauma, infecção e necrose avascular, assim como pacientes com escoliose degenerativa acentuada. Também foram excluídos casos submetidos previamente à cirurgia, radioterapia ou quimioterapia.

Dessa forma, dos 220 pacientes recuperados na busca, 63 pacientes diagnosticados com pelo menos uma fratura vertebral lombar por insuficiência foram selecionados, sendo que havia pacientes com mais de um corpo vertebral lombar fraturado. O grupo de estudo incluiu 38 mulheres e 25 homens com uma idade média de 62,25 anos. No total, foram analisadas 102 FVCs e 89 corpos vertebrais normais, sem fraturas ou outras anormalidades significativas, sendo este último grupo utilizado como controle. Os corpos vertebrais considerados normais foram extraídos da coluna lombar dos pacientes com FVCs benignas. Das 102 FVCs, 53 são fraturas por insuficiência ósseas secundárias à osteoporose e 49 são fraturas secundárias a

metástases ósseas. A Tabela 2 mostra as quantidades de corpos vertebrais lombares utilizados.

As classificações obtidas pelo software e algoritmos desenvolvidos foram comparadas com a classificação padrão-ouro baseada no diagnóstico final obtido por busca no prontuário eletrônico. Esta busca incluiu dados do RIS e do Sistema de Informação Hospitalar (HIS, em inglês "*Hospital Information System*"). Para obter a classificação padrão de referência, foram analisados relatórios de exames de diagnóstico por imagem, as imagens foram reavaliadas, assim como foram analisados os resultados de exames laboratoriais e para os casos de neoplasia maligna o resultado histopatológico obtido por meio de biópsia. A revisão do prontuário eletrônico e das imagens para esta classificação foi realizada pelo Professor Doutor Marcello Henrique Nogueira-Barbosa, radiologista com mais de 15 anos de experiência em análise de imagens musculoesqueléticas.

As metástases ósseas dos casos que compõe nossa casuística são originárias de mieloma múltiplo (12 pacientes), carcinoma mamário (oito pacientes), carcinoma pulmonar (um paciente), carcinoma prostático (dois pacientes), carcinoma de células escamosas da orofaringe (dois pacientes), colangiocarcinoma (um paciente), paraganglioma (um paciente), câncer miofibroblástico inflamatório (um paciente), leucemia linfocítica crônica (um paciente), e leucemia linfocítica aguda (um paciente).

Todos os exames foram adquiridos do equipamento de Ressonância Magnética *Philips Achieva 1.5T MRI system, Philips Medical Systems, Best, The Netherlands*. O protocolo de aquisição foi o protocolo de rotina clínica e as imagens utilizadas são as imagens do plano sagital mediano (corte central). Os parâmetros das imagens são: sequencia em ponderação T1, tempo de eco (TE) médio de $9,83 \pm 0,82\text{ms}$, tempo de repetição (TR) médio de $489 \pm 75,58\text{ms}$, campo de visão (FOV abreviatura do inglês: "*Field of view*") médio de $32,25 \times 33,33 \pm 3,98 \times 6,39\text{cm}$, matriz de aquisição média de $309,91 \times 260,57 \pm 54,84 \times 42,42$, tipo de aquisição 2D, flip angle 90° e espessura de corte de 4mm. Esse conjunto de imagens foi extraído do RIS

em formato *Digital Imaging and Communications in Medicine* (DICOM) com 16-bits/pixel (65536 níveis de cinza). Todo o conjunto foi posteriormente convertido para o formato *Tagged Image File Format* (TIFF) sem compressão e convertido pelo software desenvolvido para a escala de 8-bits/pixel (256 níveis de cinza).

Tabela 25 – Base de dados do trabalho quantificando os corpos vertebrais lombares de acordo com a classificação padrão-ouro.

Classificação	Corpos vertebrais lombares						Total FVCs
	L5	L4	L3	L2	L1	Total	
FVC maligna	9	11	10	10	9	49	102
FVC benigna	6	7	9	10	21	53	
Normal	22	21	19	18	9	89	
Total	37	39	38	38	39	191	

É importante salientar que, de acordo com o aprovado pelo CEP, as imagens foram completamente anonimizadas de forma a impedir que qualquer informação pessoal dos pacientes seja revelada para os demais pesquisadores ou para terceiros. Apenas os pesquisadores envolvidos na pesquisa tem acesso às informações do cabeçalho DICOM das imagens dos pacientes e sobre a classificação dos corpos vertebrais em normais, FVCs benignas e FVCs malignas.

5.3. Segmentação manual

Os corpos vertebrais lombares L1, L2, L3, L4 e L5 das 63 imagens selecionadas foram segmentados manualmente pelo aluno de iniciação científica Guilherme Augusto Metzner sob minha supervisão e sob a supervisão do Prof. Dr. Marcello Henrique Nogueira-Barbosa. Foi utilizado o software Adobe Photoshop (PHOTOSHOP, 2010) para segmentação manual. Essa segmentação gerou arquivos de imagens em formato TIFF sem compressão, visto na Figura 6, com o mesmo dimensionamento de linhas e colunas das imagens dos exames originais e servem como máscaras que identificam quais são os pixels dos corpos vertebrais na imagem

original. Dessa forma podemos fazer a intersecção da máscara e do exame a fim de extrair apenas os corpos vertebrais, ou seja, extrair as ROIs.

Figura 6 – Ilustração da segmentação manual de um exame de um paciente com uma FVC benigna em L1. (a) Exame de RM em ponderação T1; (b) Máscara do exame de RM gerada pela segmentação manual.

5.4. Desenvolvimento do software

Foi desenvolvido um software em MATLAB (MATLAB, 2014) com técnicas de processamento de imagens que realizam cálculos de forma, textura e análise do histograma de níveis de cinza. Ele tem como *input* as imagens dos exames e suas respectivas máscaras geradas na segmentação manual, e tem como *output* o vetor de características de cada corpo vertebral lombar, ou seja, de cada ROI. O software desenvolvido funciona em etapas segundo o fluxograma apresentado na Figura 7.

Figura 7 – Fluxograma do software de extração de características desenvolvido. O processo ocorre por 63 vezes, que é a quantidade de imagens selecionadas para o trabalho.

5.4.1. Etapa 1: Busca das imagens.

A primeira etapa do software se resume ao *input* do software que são as imagens:

- i -ésima IRM, sendo $i = 1, 2, \dots, 63$ (Figura 6a) que é convertida para 8-bits/pixel;
- Máscara correspondente à i -ésima IRM (Figura 6b) que é binarizada pelo software.

5.4.2. Etapa 2: Identificação das coordenadas dos corpos vertebrais.

Na segunda etapa, o software utiliza a máscara, como ilustrado na Figura 6b, e realiza um cálculo para identificação dos corpos vertebrais. Esse cálculo de identificação utiliza a função do MATLAB chamada “*regionprops*”. Essa função identifica a quantidade de objetos em

uma imagem e retorna quatro coordenadas (x,y) de vértices da região que envolve esses objetos, região chamada de “*Bounding Box*”, além do centroide de cada objeto. Portanto a função encontrará cinco objetos que correspondem as cinco ROIs e nos retornará suas *Bounding Boxes* e seus centroides ilustrados na Figura 8.

Figura 8 – Imagem de máscara binarizada. Os corpos vertebrais estão representados em branco, os centroides são os pontos em vermelho e as *bounding boxes* estão representadas em verde.

5.4.3. Etapa 3: Normalização dos exames

As intensidades de sinal dos tecidos humanos variam de acordo com as diferenças biofísicas do organismo de cada paciente, e também com a variação das condições de aquisição das imagens de RM, como por exemplo, o posicionamento do paciente dentro da bobina do equipamento de ressonância magnética, com as dimensões do campo de visão, entre outros fatores.

Uma das análises qualitativas utilizadas por radiologistas para a classificação entre FVCs benignas e malignas é a comparação da intensidade do sinal da IRM dos corpos vertebrais com o sinal dos discos intervertebrais. Um corpo vertebral com intensidade de sinal

da medula óssea de forma semelhante à de um disco intervertebral vizinho gera uma suspeita de malignidade (DESERNO, 2011). Podemos observar na Figura 9, apontado pelas setas, essa semelhança de intensidade de sinal.

Figura 9 – Exame de RM com uma FVC maligna em L4 secundária a metástase de carcinoma mamário. A seta branca aponta para L4 e a seta amarela aponta o disco intervertebral inferior a L4. Note neste caso a semelhança de intensidade de sinal entre o corpo vertebral e o disco intervertebral.

Desse modo, na terceira etapa, foi realizada uma normalização das imagens antes de serem processados os cálculos das características. Cada imagem é normalizada pela média dos valores de níveis de cinza dos discos intervertebrais lombares. Para esse propósito, uma matriz 5×5 pixels é posicionada nos quatro discos intervertebrais de L1 a L5 utilizando as coordenadas das *Bounding Boxes* para isso como ilustrado na Figura 10. É feito um cálculo da média dos quatro valores de média dos discos e, com essa média, a imagem é normalizada. As Equações IV e V a seguir ilustram o cálculo da normalização do exame.

$$novaImg(i,j) = \frac{imgOriginal(i,j)}{mediaDisco} \quad \dots\dots\dots (9)$$

$$ImgNormalizada(i,j) = 255 \times \frac{(novaImg(i,j) - \min(novaImg))}{\max(novaImg) - \min(novaImg)} \quad \dots\dots\dots (10)$$

Figura 10 – Os contornos dos corpos vertebrais estão mostrados nas linhas verdes. As bounding boxes de cada corpo vertebral estão mostradas nas linhas amarelas. Em azul são regiões dos discos intervertebrais gerados automaticamente utilizando as bounding boxes dos corpos vertebrais como referência. Em vermelho estão quatro blocos de 5×5 pixels cada, automaticamente posicionados nos discos intervertebrais entre L1-L2, L2-L3, L3-L4 e L4-L5 utilizando as bounding boxes vertebrais como referência.

5.4.4. Etapa 4: Intersecção dos exames em RM e suas máscaras

Na quarta etapa, o software realiza a intersecção entre a imagem normalizada e sua imagem máscara resultando em uma terceira imagem ilustrada na Figura 11. Portanto os pixels da imagem normalizada que não pertencerem às cinco ROIs, na nova imagem terão valor igual a 0, caracterizado pelo fundo preto da imagem.

Figura 11 – Exemplo de imagem resultante da Etapa 4 do software.

5.4.5. Etapa 5: Detecção das ROIs.

Na quinta etapa, o software utiliza a imagem resultante da intersecção com as coordenadas das *Bounding Boxes*, gera cinco imagens com o conteúdo de cada ROI.

Há casos onde determinada *Bounding Box de um corpo vertebral* envolve uma parte do corpo vertebral adjacente devido à curvatura da coluna vertebral paciente como ilustrado na Figura 12. Essa parte que não pertence à *Bounding Box* do corpo vertebral em questão é considerada um ruído para a ROI. A fim de detectar essa região de ruído na *bounding box*, foi aplicado um algoritmo recursivo chamado *Flood Fill* com oito direções (SOILLE, 2013) que determina as áreas formadas por pixels relacionados que estão conectados a uma distância de um pixel. Esse algoritmo detecta os pixels pertencentes tanto ao corpo vertebral quanto ao ruído e este foi removido com o comando $ruído(i,j) = 0$ como ilustrado na Figura 13.

A Figura 14 ilustra nove ROIs preparadas nessa etapa do software sendo três exemplos de cada classe em estudo: normal, FVC benigna e FVC maligna.

Figura 12 – Exemplo de duas ROIs e suas *bounding boxes* em amarelo. A seta branca indica o ruído do corpo vertebral adjacente envolvido pela *bounding box* que é removido pelo software.

Figura 13 – *Bounding box* com ruído apresentada na Figura 12. (a) a seta branca aponta para o ruído na imagem; (b) a seta branca destaca que o ruído foi removido.

Figura 14 – Exemplos de imagens de ROIs dos corpos vertebrais geradas na Etapa 5 do software. Através dessas ROIs são extraídas as características dos corpos vertebrais. A linha (a) representa três exemplos de ROIs com corpos vertebrais com padrão de normalidade; a linha (b) representa três exemplos de ROIs com FVCs de padrão benigno; a linha (c) representa três exemplos de ROIs com FVCs de padrão maligno.

5.4.6. Etapa 6: Cálculo das características das ROIs.

A sexta etapa do software é a etapa onde ocorre o cálculo das características, descritores ou atributos de forma, textura e níveis de cinza dos corpos vertebrais, presentes nas ROIs geradas na Etapa 5, os quais irão compor seus vetores características na classificação.

5.4.6.1. Atributos de níveis de cinza

O software extrai três atributos estatísticos de níveis de cinza dos corpos vertebrais calculados com valores dos seus histogramas, lembrando que as ROIs estão em escala de

cinza de 8-*bits*/pixel. Portanto o histograma é representado computacionalmente por um arranjo ou vetor com 256 índices representados nas equações abaixo como $x_1, x_2, \dots, x_i, \dots, x_{256}$ e ilustrado anteriormente na Figura 3. Os atributos calculados são:

i. Coeficiente de Variação (cv):

$$cv = \frac{\sigma}{\mu} \quad \text{..... (11)}$$

Onde σ , e média, μ , dos níveis de cinza do histograma:

$$\mu = \frac{1}{255} \sum_{i=1}^{256} x_i \quad \text{..... (12)}$$

$$\sigma = \sqrt{\frac{1}{255} \sum_{i=1}^{256} ((x_i - \mu)^2)} \quad \text{..... (13)}$$

ii. Skewness ($skew$):

Skewness é uma medida de assimetria do histograma (JOANES e GILL, 1998). Se $skew = 0$, significa que o histograma possui uma distribuição simétrica. Se $skew < 0$, isso indica que sua distribuição está mais deslocada à esquerda, ou seja, maior quantidade de pixels mais escuros. Porém se $skew > 0$, o histograma terá uma distribuição dos níveis de cinza mais a direita, ou seja, maior quantidade de pixels com níveis de cinza mais claros.

$$skew = \frac{1}{255 \times \sigma^3} \sum_{i=1}^{256} (x_i - \mu)^3 \quad \text{..... (14)}$$

iii. Kurtosis (*kurt*):

Kurtosis é uma medida de achatamento ou pico dos dados do histograma em relação à curva normal ou curva de Gauss (JOANES e GILL, 1998). Um histograma distribuído em uma curva de Gauss possui $kurt = 0$ e possui uma função mesocúrtica. Um valor de $kurt > 0$ significa que a distribuição está formando picos e, portanto é mais alta que a curva de Gauss. Um valor de $kurt < 0$ significa que a distribuição do histograma é mais achatada que a curva de Gauss, portanto sua curva é mais baixa.

$$kurt = \left(\frac{1}{255 \times \sigma^4} \sum_{i=1}^{256} (x_i - \mu)^4 \right) - 3 \quad \dots\dots\dots(15)$$

5.4.6.2. Atributos de textura

Como análise de textura, são utilizados 14 descritores de Haralick et al. (1973). Trata-se de uma análise estatística da distribuição dos valores de pixels e a relação entre eles. Esse inter-relacionamento entre os níveis de cinza é representado pela matriz de coocorrência dos níveis de cinza, *MC*. Assim como realizado para os atributos de níveis de cinza, apenas os pixels internos ao corpo vertebral foram utilizados para o cálculo, ignorando o fundo preto da ROI. A intenção da conversão imagem de 16-bits/pixel para 8-bits/pixel é de que o cálculo das matrizes de coocorrência não exceda a capacidade de memória computacional. O software calcula a matriz de coocorrência com relação aos quatro ângulos $\theta = 0^\circ, 45^\circ, 90^\circ$ e 135° e a uma distância $d = 1$. *MC* é normalizada, *p*, calculando a soma dos elementos das colunas da matriz de coocorrência resultando em uma matriz 1×256 e, posteriormente, a soma dos elementos dessa matriz resultante. O resultado desse somatório resulta em um número real *rs*, o qual dividirá todos os valores da matriz de coocorrência. Portanto a matriz de coocorrência normalizada $p = MC/rs$. Com *p* são calculados os 14 atributos de textura:

- i. Segundo momento angular (Energia):

$$h_1 = \sum_i \sum_j \{p(i, j)\}^2 \quad \dots\dots\dots (16)$$

- ii. Contraste

$$h_2 = \sum_{n=0}^{256-1} n^2 \left\{ \sum_{i=1}^{256} \sum_{j=1}^{256} p(i, j) \right\}_{|i-j|=n} \quad \dots\dots\dots (17)$$

- iii. Correlação

$$h_3 = \frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad \dots\dots\dots (18)$$

Onde $\mu_x, \mu_y, \sigma_x, \sigma_y$ são as médias e desvios padrões de p_x e p_y .

Sendo

$$p_x(i) = \sum_{j=1}^{Ng} p(i, j) \quad \dots\dots\dots (19)$$

$$p_y(j) = \sum_{i=1}^{Ng} p(i, j) \quad \dots\dots\dots (20)$$

- iv. Soma de Quadrados: Variância

$$h_4 = \sum_i \sum_j (i - \mu)^2 p(i, j) \quad \dots\dots\dots (21)$$

- v. Diferença Inversa de Momento

$$h_5 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i, j) \quad \dots\dots\dots (22)$$

vi. Média da Soma

$$h_6 = \sum_{i=2}^{2 \times 256} i p_{x+y}(i) \quad \dots\dots\dots (23)$$

Onde

$$p_{x+y}(k) = \sum_{i=1}^{256} \sum_{j=1}^{256} p(i, j) \quad \dots\dots\dots (24)$$

Sendo $k = i + j$ $k = 2, 3, \dots, 2 \times 256$

vii. Variância da Soma

$$h_7 = \sum_{i=2}^{2 \times 256} (i - h_6)^2 p_{x+y}(i) \quad \dots\dots\dots (25)$$

viii. Entropia da Soma

$$h_8 = - \sum_{i=2}^{2 \times 256} p_{x+y}(i) \log\{p_{x+y}(i)\} \quad \dots\dots\dots (26)$$

ix. Entropia

$$h_9 = - \sum_i \sum_j p(i, j) \log(p(i, j)) \quad \dots\dots\dots (27)$$

x. Diferença da Variância

$$h_{10} = \text{variância de } p_{x-y} \quad \dots\dots\dots (28)$$

xi. Diferença da Entropia

$$h_{11} = - \sum_{i=0}^{256-1} p_{x-y}(i) \log\{p_{x-y}(i)\} \quad \dots\dots\dots (29)$$

Sendo:

$$p_{x-y}(k) = \sum_{i=1}^{256} \sum_{j=1}^{256} p(i,j) \quad \dots\dots\dots (30)$$

$$k = |i - j| \quad k = 0, 1, \dots, 256 - 1$$

xii. Medidas de informação de correlação

$$h_{12} = \frac{HXY - HXY1}{\max\{HX, HY\}} \quad \dots\dots\dots (31)$$

$$h_{13} = (1 - \exp[-2.0(HXY2 - HXY)])^{1/2} \quad \dots\dots\dots (32)$$

$$HXY = - \sum_i \sum_j p(i,j) \log(p(i,j)) \quad \dots\dots\dots (33)$$

HX e HY são entropia de p_x e p_y .

$$HXY1 = - \sum_i \sum_j p(i,j) \log\{p_x(i)p_y(i)\} \quad \dots\dots\dots (34)$$

$$HXY2 = - \sum_i \sum_j p_x(i)p_y(i) \log\{p_x(i)p_y(i)\} \quad \dots\dots\dots (35)$$

xiii. Coeficiente de máxima correlação

$$h_{14} = (\text{segundo maior autovalor de } Q)^{1/2} \quad \dots\dots\dots (36)$$

Onde

$$Q(i,j) = \sum_k \frac{p(i,j)p(j,k)}{p_x(i)p_y(k)} \quad \dots\dots\dots (37)$$

5.4.6.3. Atributos de forma

Como atributos de forma, foram extraídos 38 atributos. Para cálculos de forma dos corpos vertebrais, foram utilizadas ROIs geradas a partir das imagens das máscaras binarizadas na Etapa 5 como ilustrado na Figura 15.

Figura 15 – Mesmas ROIs dos corpos vertebrais ilustrados na Figura 14 após a binarização para a utilização nos cálculos de atributos de forma. A linha (a) representa três exemplos de ROIs com corpos vertebrais com padrão de normalidade; a linha (b) representa três exemplos de ROIs com FVCs de padrão benigno; a linha (c) representa três exemplos de ROIs com FVCs de padrão maligno.

i. Compacidade (C_o)

Compacidade é definida como $C = P^2/V_{area}$, onde P é o perímetro do contorno do corpo vertebral e V_{area} é a área contida no interior do corpo vertebral. Essa medida é invariante a translação, rotação, escala e reflexão de um contorno, como desejado nas aplicações de análise de forma. Porém o valor de C aumenta com o alongamento, entalhe e a rugosidade da forma do corpo vertebral. Uma versão normalizada de C foi computada de acordo com Rangayyan (2005):

$$C_o = 1 - \frac{4\pi V_{area}}{P^2} \quad \text{..... (38)}$$

onde C_o varia de $[0,1]$, sendo $C_o = 0$ um círculo. Esse atributo foi estudado em aplicações de análise de câncer de pulmão (CABRAL e RANGAYYAN, 2012; RAGUSO *et al.*, 2010). A Figura 16 ilustra P e V_{area} de uma ROI.

Figura 16 – Representação do contorno do corpo vertebral na cor vermelha onde o perímetro P é calculado e na cor branca onde a área vertebral V_{area} foi calculada.

ii. Deficiência convexa (CD)

A diferença entre as áreas contidas pelo contorno do corpo vertebral (V_{area}) e seu casco convexo (em inglês, *convex hull*) (CH_{area}) dividido por V_{area} define a deficiência convexa do corpo vertebral. CH_{area} portanto é a área do polígono circunscrito ao corpo vertebral:

$$CD = \frac{V_{area} - CH_{area}}{V_{area}} \quad \dots\dots\dots (39)$$

Se $CD = 0$, o corpo vertebral seria um objeto convexo comparável a um retângulo, ou seja, quanto menor o valor de CD menor a concavidade vertebral. A Figura 17 ilustra os parâmetros utilizados no cálculo de CD .

Figura 17 – O polígono em verde representa o *convex hull* do corpo vertebral e em vermelho está o contorno do corpo vertebral.

iii. Descritor de Fourier (FDF)

O descritor de Fourier (GRANLUND, 1972; ZAHN e ROSKIES, 1972; PERSON e FU, 1977; RANGAYYAN, 2005) é definido como:

$$Z(k) = \frac{1}{N} \sum_{n=0}^{N-1} z(n) \exp \left[-j \frac{2\pi}{N} nk \right] \quad \dots\dots\dots (40)$$

onde $k = -N/2 + 1, \dots, -1, 0, 1, 2, \dots, N/2$, $z(n) = x(n) + jy(n)$, $n = 0, 1, \dots, N - 1$, são coordenadas de N pixels do contorno (RANGAYYAN, 2005). Uma vez que a mesma sequência de pixels de contorno é obtida, se o contorno for repetidamente atravessado, $z(n)$ é um sinal periódico. Consequentemente, $Z(k)$ representa uma série de Fourier com k sendo o índice de harmônicos de uma frequência fundamental. $Z(1)$ e $Z(-1)$ representam círculos centralizados nas coordenadas dadas por $Z(0)$. O uso de $Z(-1)$, $Z(0)$ e $Z(1)$ juntos na equação inversa de Fourier leva a elipse de melhor ajuste ao contorno dado.

Contornos rugosos causam maiores valores de $Z(k)$ para altos valores de k do que para contornos suaves. A fim de caracterizar as mudanças de forma causadas por FVCs, em particular a rugosidade causada por FVCs benignas e as mudanças mais suaves causadas por FVCs malignas, foi definido o atributo descritor de Fourier (FDF) como:

$$FDF = \frac{\sum_{k=k_1}^{N/2} |Z(k)|^2 + \sum_{k=-N/2+1}^{-k_1} |Z(k)|^2}{\sum_{k=-N/2+1}^{N/2} |Z(k)|^2} \quad \dots\dots\dots (41)$$

com $Z(0) = 0$. FDF representa a fração do total de energia presente na harmônica de maior ordem que $\pm k_1$, não incluindo $Z(0)$. Depois de conduzir várias seleções de atributos e experimentos de classificação com vários valores de k_1 , seu valor foi definido nesse trabalho como $k_1 = \lfloor N/12 \rfloor$. FDF é normalizado para a escala de $[0,1]$, com valores altos para formas mais rugosas, correspondendo a maiores valores de FDF em maiores frequências.

iv. Momentos centrais invariantes de HU (M_i)

Hu (HU, 1962) propôs sete fatores de forma, definidos nesse trabalho como M_1 a M_7 , baseados nos momentos de figuras geométricas planas, os quais foram usados em outras aplicações de análise de forma (RANGAYYAN, 2005). Esses atributos, da mesma forma que

são efetivos na classificação de vários tipos de formas, são de difícil interpretação e tem relação com variações de forma em uma aplicação específica. Os momentos de Hu são definidos como:

$$M_1 = \mu_{20} + \mu_{02} \quad \text{..... (42)}$$

$$M_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \quad \text{..... (43)}$$

$$M_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \quad \text{..... (44)}$$

$$M_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \quad \text{..... (45)}$$

$$M_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \\ + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \quad \text{..... (46)}$$

$$M_6 = (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] + 4\mu_{11}(\mu_{30} + \mu_{12})(\mu_{21} + \mu_{03}) \quad \text{..... (47)}$$

$$M_7 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] \\ - (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \quad \text{..... (48)}$$

Onde μ_{pq} são os momentos centrais calculados como:

$$\mu_{00} = m_{00} = \mu \quad \text{..... (49)}$$

$$\mu_{10} = \mu_{01} = 0 \quad \text{..... (50)}$$

$$\mu_{20} = m_{20} - \mu\bar{x}^2 \quad \text{..... (51)}$$

$$\mu_{11} = m_{11} - \mu\bar{x}\bar{y} \quad \text{..... (52)}$$

$$\mu_{02} = m_{02} - \mu\bar{y}^2 \quad \text{..... (53)}$$

$$\mu_{30} = m_{30} - 3m_{20}\bar{x} + 2\mu\bar{x}^3 \quad \dots\dots\dots (54)$$

$$\mu_{21} = m_{21} - m_{20}\bar{y} - 2m_{11}\bar{x} + 2\mu\bar{x}^2\bar{y} \quad \dots\dots\dots (55)$$

$$\mu_{12} = m_{12} - m_{02}\bar{x} - 2m_{11}\bar{y} + 2\mu\bar{x}\bar{y}^2 \quad \dots\dots\dots (56)$$

Onde m_{pq} são os momentos calculados como:

$$m_{pq} = \sum_i \sum_j i^p j^q \text{img}(i, j) \quad , \quad p, q = 0, 1, 2, \dots \quad \dots\dots\dots (57)$$

$$\bar{x} = \frac{m_{10}}{m_{00}} \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

v. Atributos estatísticos das distâncias vertebrais ao centroide

As distâncias Euclidianas do centroide da ROI até cada pixel do contorno do corpo vertebral foi calculado conforme ilustrado na Figura 18. A unidade de medida dessas distâncias está em milímetros (mm). Sete atributos estatísticos foram calculados para cada ROI: a média (*meanDCB*), desvio padrão (*stdDCB*), variância (*varDCB*), coeficiente de variação (*cvDCB*), *skewness* (*skewDCB*), *kurtosis* (*kurtDCB*), e a combinação dos momentos de quarta e segunda ordem (*mf*) como definido por Shen et al. (1994):

$$mf = \frac{\left\{ \frac{1}{N} \sum_{n=0}^{N-1} [d(n) - md]^4 \right\}^{1/4} - \left\{ \frac{1}{N} \sum_{n=0}^{N-1} [d(n) - md]^2 \right\}^{1/2}}{md} \quad \dots\dots\dots (58)$$

onde md é a média das distâncias euclidianas $d(n)$, $n = 0$ a $N - 1$.

Figura 18 – Em ciano estão ilustradas as distâncias entre o contorno dos corpos vertebral e seus centroides em vermelho. (a) corpo vertebral dentro da normalidade; (b) FVC benigna; (c) FVC maligna.

vi. Atributos estatísticos das alturas e larguras vertebrais

Os corpos vertebrais presentes nas ROIs possuem inclinações diferentes, por isso faz-se necessário o uso de alguma técnica matemática para calcular suas alturas e larguras.

Os momentos espaciais de uma imagem podem ser usados para determinar o eixo principal de objetos presentes na imagem, o qual pode ser útil para encontrar o ângulo dominante de alinhamento direcional. O momento de inércia está em seu mínimo quando o momento é tomado em torno do eixo principal e o centroide da imagem. A direção do eixo principal, θ , é dada por (RANGAYYAN, 2005; ROSENFELD e KAK, 2014):

$$\tan 2\theta = \frac{2\mu_{11}}{(\mu_{20} - \mu_{02})} \quad \dots\dots\dots (52)$$

Onde μ_{11} , μ_{20} , e μ_{02} são os momentos centrais da imagem como calculados anteriormente para os momentos de Hu.

A linha que passa através do centroide do corpo vertebral com o ângulo θ é o seu eixo principal (RANGAYYAN, 2005; ROSENFELD e KAK, 2014). Com θ foi possível calcular linhas paralelas e perpendiculares ao eixo principal em toda a extensão do corpo vertebral, as quais

representam as alturas e larguras. Portanto valores de altura na direção superior-inferior (h_{si}) e de largura na direção anterior-posterior (w_{ap}) foram calculados utilizando distância Euclidiana. A Figura 19 ilustra exemplos de ROIs com seus eixos principais e suas linhas de alturas e larguras vertebrais calculadas em relação aos eixos principais.

Figura 19 – As linhas em vermelho representam o eixo principal das ROIs, as linhas azuis perpendiculares ao eixo principal representam as alturas superior-inferior e as linhas verdes paralelas ao eixo principal representam as larguras anterior-posterior. (a) ROI de um corpo vertebral normal, (b) ROI de um corpo vertebral com uma FVC padrão benigno e (c) ROI de um corpo vertebral com uma FVC padrão maligno.

As medidas de largura e altura foram escaladas pelos parâmetros de imagem relacionados para obter os comprimentos correspondentes em mm. Para cada conjunto de medidas h_{si} , w_{ap} e da união de ambas, de um corpo vertebral foram calculados atributos estatísticos presentes na Tabela 3.

Tabela 26 – Atributos estatísticos calculados a partir das alturas superior-inferior, das larguras anterior-posterior e da união de ambas.

Medida estatística	h_{si}	w_{ap}	$h_{si} \cup w_{ap}$
Média	<i>meanLineV</i>	<i>meanLineH</i>	<i>meanLineHV</i>
Desvio Padrão	<i>stdLineV</i>	<i>stdLineH</i>	<i>stdLineHV</i>
Variância	<i>varLineV</i>	<i>varLineH</i>	<i>varLineHV</i>
Coeficiente de Variação	<i>cvLineV</i>	<i>cvLineH</i>	<i>cvLineHV</i>
Skewness	<i>skewLineV</i>	<i>skewLineH</i>	<i>skewLineHV</i>
Kurtosis	<i>kurtLineV</i>	<i>kurtLineH</i>	<i>kurtLineHV</i>

Além disso, as medidas de h_{si} foram divididas em três subconjuntos iguais correspondentes às porções anterior, central e posterior do corpo vertebral, e as médias de cada subconjunto foram calculadas, sendo M_A , M_C , e M_P respectivamente. Três atributos foram calculados através da diferença dessas médias: $\mu AC = M_A - M_C$, $\mu CP = M_C - M_P$ e $\mu PA = M_P - M_A$. No total, 21 atributos de forma foram extraídos a partir da altura e largura vertebral.

5.4.7. Etapa 7: Organização dos vetores de características

No total, foram calculados 55 atributos, sendo três de níveis de cinza, 14 de textura e 38 de forma. Cada ROI terá o seu vetor de características composto pelos 55 atributos. A seguir está sumarizada a relação dos 55 atributos calculados na Etapa 6 presentes no vetor de características.

i. Níveis de cinza:

- cv
- $skew$
- $kurt$

ii. Textura:

- h_1
- h_2
- h_3
- h_4
- h_5
- h_6
- h_7
- h_8
- h_9
- h_{10}
- h_{11}
- h_{12}
- h_{13}
- h_{14}

iii. Forma:

• C_o	• $cvLineV$	• $skewLineHV$
• FDF	• $skewLineV$	• $kurtLineHV$
• CD	• $kurtLineV$	• μAC
• M_1	• $meanLineH$	• μCP
• M_2	• $stdLineH$	• μPA
• M_3	• $varLineH$	• $meanDCB$
• M_4	• $cvLineH$	• $stdDCB$
• M_5	• $skewLineH$	• $varDCB$
• M_6	• $kurtLineH$	• $cvDCB$
• M_7	• $meanLineHV$	• $skewDCB$
• $meanLineV$	• $stdLineHV$	• $kurtDCB$
• $stdLineV$	• $varLineHV$	• mf
• $varLineV$	• $cvLineHV$	

5.4.8. Etapa 8: Arquivos (*output*).

Essa é a etapa onde os cinco vetores características do exame de RM correspondentes as cinco ROIs dos corpos vertebrais lombares, são escritos em seus respectivos arquivos em formato TXT: L1.txt, L2.txt, L3.txt, L4.txt e L5.txt.

Ao fim das oito etapas para uma imagem, o software retorna à primeira etapa a fim de buscar uma nova imagem e todo o processo se repete para as 63 imagens.

5.5. Classificação e estatística

Essa é a etapa do trabalho onde foi realizado o treinamento do computador para a detecção e classificação de FVCs. Etapa que envolveu o uso de algoritmos de inteligência artificial e aprendizado de máquinas, além de análises estatísticas.

A base de dados possui corpos vertebrais com três classes: FVC maligna, FVC benigna e normal. Cada vetor de características extraído pelo software desenvolvido recebeu manualmente sua respectiva classe com o objetivo de se realizar o treinamento dos algoritmos classificadores. Ou seja, ao final de todo vetor de características foi inserido o atributo “classe”. A partir disso foram realizadas três análises de classificação independentes entre si:

i. FVC x Normal

Análise de classificação binária a qual a classe “FVC” engloba as FVCs benignas e FVCs malignas e será comparada com a classe “normal” onde estão os casos de corpos vertebrais sem fratura e dentro dos padrões de normalidade.

Figura 20 – Ilustração da classificação de corpos vertebrais com FVC x Normal.

ii. FVC Maligna x FVC benigna

Assim como a análise anterior, trata-se de uma análise de classificação binária a qual envolve apenas os casos de FVCs.

Figura 21 – Ilustração da classificação de corpos vertebrais com FVC Maligna x FVC Benigna.

iii. FVC Maligna x FVC benigna x Normal

Diferentemente das análises de classificação anteriores, essa realiza a classificação com as três classes.

Figura 22 – Ilustração da classificação de corpos vertebrais com FVC Maligna x FVC Benigna x Normal.

5.5.1. Estatística de cada atributo

Análises estatísticas foram realizadas para cada atributo dos vetores de características para as duas análises de classificação binárias: FVC x Normal e FVC Maligna x FVC Benigna. O teste de Kolmogorov-Smirnov (MASSEY JR, 1951) foi aplicado para checar a normalidade das distribuições dos atributos. Como todos os atributos em ambas as classificações rejeitaram a hipótese nula de que as amostras seguem uma distribuição normal, foi utilizado o teste de Wilcoxon-Mann-Whitney, também chamado de Wilcoxon rank-sum ou teste U (WILCOXON et

al., 1970) para computar os p-valores para analisar significatividade estatística de cada atributo em cada análise de classificação. Além disso, a estimada A_z ou AUROC de cada atributo em cada análise de classificação foi calculada usando o software ROCKIT versão 0.9 (METZ, 1998).

5.5.2. Seleção de atributos

Os atributos desenvolvidos nesse trabalho passaram por uma série de estudos e discussões antes de serem adicionados ao vetor de características, porém, a fim de reduzir a quantidade de possíveis atributos irrelevantes, foram realizadas seleções de atributos nos vetores de características antes das classificações.

i. Todos os atributos

Para cada classificador utilizado nas três análises do trabalho (FVC x Normal; FVC Maligna x FVC Benigna; FVC Maligna x FVC Benigna x Normal), foi realizada a classificação sem a seleção de atributos.

ii. Atributos com $p < 0,001$ e cinco maiores valores de A_z ($5+ A_z$)

Foram escolhidos para as análises de classificações FVC x Normal e FVC Maligna x FVC Benigna os atributos que apresentaram forte significância estatística ($p < 0,001$) e os cinco maiores valores de A_z para cada uma das análises. Portanto o vetor de características após essa seleção de atributos passa a ter cinco atributos e o mesmo vetor de características é utilizado em todos os classificadores utilizados.

iii. Atributos com $p < 0,001$ e dez maiores valores de A_z ($10 + A_z$)

Foram utilizados para as análises de classificações FVC x Normal e FVC Maligna x FVC Benigna os atributos com $p < 0,001$ e os dez maiores valores de A_z para cada uma das análises. Portanto o vetor de características após essa seleção de atributos passa a ter dez atributos e o mesmo vetor de características é utilizado em todos os classificadores utilizados.

iv. Método automático *Wrapper*

A seleção de atributos automática utilizando o método *Wrapper* foi realizada no software de aprendizado de máquinas WEKA (HALL *et al.*, 2009). Esse método está presente na aba “*Select Attributes*” com o nome “*WrapperSubsetEval*”.

Para cada classificador utilizado nas três análises do trabalho (FVC x Normal; FVC Maligna x FVC Benigna; FVC Maligna x FVC Benigna x Normal), o método *Wrapper* de seleção de atributos foi executado com os seguintes parâmetros: número de *folds* = 10, *seed* = 1 e *threshold* = 0,01. Foi utilizado o algoritmo *best-first* como método de busca desses subconjuntos com os seguintes parâmetros: direção *forward*, *lookupCacheSize* = 1, *searchTermination* = 5 e *startSet* = *null*. A *ACC* foi estimada utilizando o modo de seleção de atributos em *cross-validation* com os parâmetros: número de *folds* = 10 e *seed* = 1.

A resposta dada pela seleção de atributos são as porcentagens as quais cada atributo foi selecionado em 10 *folds*. Dessa forma, foram selecionados para compor o novo vetor de características os atributos selecionados iguais ou acima de 30% e, portanto, cada vetor característica gerado é específico de cada classificador utilizado.

5.5.3. Algoritmos de classificação e validação dos resultados

As classificações foram realizadas utilizando algoritmos presentes no pacote do software WEKA (HALL *et al.*, 2009). Todas as classificações foram feitas na aba do software

com o nome “*Classify*” utilizando a técnica de amostragem *cross-validation* estratificada com a quantidade de *folds* = 10.

Os resultados da classificação foram calculados através da matriz de confusão de melhor custo benefício fornecida pelo software WEKA. A validação dos resultados foi feita em cima dos valores de sensibilidade, especificidade, acurácia, calculados pela matriz de confusão ótima exibida nos resultados do WEKA, e média ponderada da AUROC para as análises “FVC x Normal” e “FVC Maligna x FVC Benigna”; para a análise com as três classes “FVC Maligna x FVC Benigna x Normal”, o resultado foi validado em cima dos valores de acurácia e média ponderada da AUROC.

A Figura 23 ilustra o fluxo dos vetores de características quanto à seleção de atributos e a classificação.

Figura 23 – Fluxograma da seleção e classificação dos vetores de características: **a** representa o vetor de características com todos os atributos que é utilizado diretamente para a classificação em todos os classificadores; **b** representa o vetor de características com todos os atributos que é submetido ao método automático *Wrapper* de seleção de atributos para cada classificador em específico resultando em **b₁, b₂, b₃, b₄, b₅, b₆, b₇, b₈**; **c** representa o vetor de características com todos os atributos que é submetido ao método seleção de atributos 5+ A_z resultando em **c₁** que é utilizado para a classificação em todos os classificadores; **d** representa o vetor de características com todos os atributos que é submetido ao método seleção de atributos 10+ A_z resultando em **d₁** que é utilizado para a classificação em todos os classificadores.

i. Árvore de decisão J48

Os valores dos parâmetros utilizados na execução do algoritmo J48 (caminho do pacote: *trees* → *J48*) no software WEKA foram: *binarySplits* = *false*; *confidenceFactor* = 0,25; *debug* = *false*; *minNumObj* = 2; *numFolds* = 3; *reducedErrorPruning* = *false*; *saveInstanceData* = *false*; *seed* = 1; *subtreeRaising* = *true*; *unpruned* = *false*; *useLaplace* = *false*.

ii. *k*-nearest neighbor (*k*-NN)

Os valores dos parâmetros utilizados na execução do algoritmo *k*-NN (caminho do pacote: *lazy* → *IBk*) no software WEKA foram: *k* = 5, 7, 9 e 11; *crossValidate* = *true*; *debug* = *false*; *distanceWeighting* = *No distance weighting*; *meanSquared* = *false*; *nearestNeighbourSearchAlgorithm* = *LinearNNSearch (EuclidianDistance)*; *windowSize* = 0.

iii. Naïve Bayes

Os valores dos parâmetros utilizados na execução do algoritmo naïve Bayes (caminho do pacote: *bayes* → *NaïveBayes*) no software WEKA foram: *debug* = *false*; *displayModelInOldFormat* = *false*; *useKernelEstimator* = *false*; *useSupervisedDiscretization* = *false*.

iv. Radial basis function network (RBF network)

Os valores dos parâmetros utilizados na execução do algoritmo RBF *network* (caminho do pacote: *functions* → *RBFNetwork*) no software WEKA foram: *clusteringSeed* = 1; *debug* = *false*; *maxIts* = -1; *minStdDev* = 0,1; *numClusters* = 2; *ridge* = 1,0E-8.

v. Support vector machine (SVM)

Os valores dos parâmetros utilizados na execução do algoritmo SVM (caminho do pacote: *function* → *LibSVM*) no software WEKA foram: *SVMType* = *C-SVC (classification)*;

cacheSize = 40,0; *coef0* = 0,0; *cost* = 1,0; *debug* = *false*; *degree* = 3; *doNotReplaceMissingValues* = *false*; *eps* = 0,001; *gamma* = 0; *kernelType* = *linear: (u*v)*; *loss* = 0,1; *normalize* = *false*; *nu* = 0,5; *probabilityEstimates* = *false*; *seed* = 1; *shrinking* = *true*; *weights* = *null*.

5.5.4. Classificação segundo radiologistas e comparação com classificadores

Foi preparada uma apresentação de slides contendo cada uma das ROIs avaliadas pelo sistema computadorizado apresentadas em ordem aleatória, sendo que, em cada slide, o radiologista pode avaliar uma ROI. Essas ROIs foram analisadas por três radiologistas de forma independente, ou seja, os radiologistas não tiveram contato ou ajuda de outro radiologista, e também não tinham informações a respeito dos casos ou sobre leituras previamente realizadas:

- i. Radiologista 1: Médico residente em fase de subespecialização em radiologia musculoesquelética, no quarto ano de residência médica (R4)
- ii. Radiologista 2: Médico residente de Radiologia, no terceiro ano de residência médica (R3)
- iii. Radiologista 3: radiologista geral com 14 anos de experiência.

Os radiologistas classificaram as ROIs como FVC maligna, FVC benigna e Normal, e a partir disso foram geradas matrizes de confusão para as mesmas análises realizadas para os classificadores:

i. FVC x Normal

Classificações dadas como “FVC Maligna” e “FVC Benigna” foram consideradas de classe “FVC” nessa análise.

ii. FVC Maligna x FVC Benigna

Classificações errôneas de FVC Maligna ou FVC Benigna para casos da classe “Normal” não entraram na contagem da matriz de confusão.

Foram comparados valores de acurácia, sensibilidade e especificidade com intervalo de confiança de 95% dos radiologistas e dos classificadores com atributos selecionados a partir do método de seleção de atributo cuja média aritmética de AUROC das classificações seja a maior.

6. RESULTADOS

6.1. Estatística dos atributos: p-valores e A_z

i. FVC x Normal:

A Tabela 4 traz os resultados dos p-valores calculados pelo teste Wilcoxon-Mann-Whitney e os valores de A_z calculados pelo software ROCKIT versão 0.9 para cada atributo na análise entre FVC x Normal.

ii. FVC Maligna x FVC Benigna:

A Tabela 5 traz os resultados dos p-valores calculados pelo teste Wilcoxon-Mann-Whitney e os valores de A_z calculados pelo software ROCKIT versão 0.9 para cada atributo na análise entre FVC Maligna x FVC Benigna.

Tabela 27 – p-valores calculados pelo teste Wilcoxon-Mann-Whitney e os valores de A_z para cada atributo na análise FVC x Normal. Cores das condições: $p < 0,001$ e $A_z \geq 0,9$ (preenchimento cinza escuro e escrito em branco); $0,001 \leq p \leq 0,01$ e $0,8 \leq A_z \leq 0,89$ (preenchimento cinza intermediário e escrito em preto); $0,01 \leq p \leq 0,05$ e $0,7 \leq A_z \leq 0,79$ (preenchimento cinza claro e escrito em preto); $p > 0,05$ e $A_z < 0,7$ ou valor desconhecido (?) (sem preenchimento de cor e escrito em preto).

FVC x Normal			
	Atributos	p - valor	A_z
Níveis de Cinza	<i>cv</i>	0,00000	0,7507
	<i>skew</i>	0,39016	0,5492
	<i>kurt</i>	0,42736	0,5317
Textura	<i>h₁</i>	0,01060	0,4218
	<i>h₂</i>	0,00032	0,661
	<i>h₃</i>	0,00328	0,6293
	<i>h₄</i>	0,00623	0,6417
	<i>h₅</i>	0,04571	0,5773
	<i>h₆</i>	0,00000	0,7307
	<i>h₇</i>	0,00633	0,6403
	<i>h₈</i>	0,00680	0,6204
	<i>h₉</i>	0,00291	0,6471
	<i>h₁₀</i>	0,00003	0,6737
	<i>h₁₁</i>	0,00277	0,6324
	<i>h₁₂</i>	0,53145	0,5241
	<i>h₁₃</i>	0,01044	0,6142
	<i>h₁₄</i>	0,41674	0,4624
	<i>C_o</i>	0,00000	0,864
	<i>FDF</i>	0,19309	0,4492
	<i>CD</i>	0,00000	0,8813
	<i>M₁</i>	0,00000	0,9636
	<i>M₂</i>	0,00000	0,9315
Forma	<i>M₃</i>	0,00000	0,8866
	<i>M₄</i>	0,00000	0,936
	<i>M₅</i>	0,00938	?
	<i>M₆</i>	0,59136	0,4975
	<i>M₇</i>	0,21174	?
	<i>meanLineV</i>	0,00000	0,9019
	<i>stdLineV</i>	0,00000	0,8615
	<i>varLineV</i>	0,00000	0,8615
	<i>cvLineV</i>	0,27561	0,5435
	<i>skewLineV</i>	0,00000	0,8569
	<i>kurtLineV</i>	0,00147	0,6294
	<i>meanLineH</i>	0,00000	0,2976
	<i>stdLineH</i>	0,00000	0,7306
	<i>varLineH</i>	0,00000	0,7306
	<i>cvLineH</i>	0,00000	0,8607
	<i>skewLineH</i>	0,00000	0,894
	<i>kurtLineH</i>	0,00000	0,8491
	<i>meanLineHV</i>	0,00000	0,8474
	<i>stdLineHV</i>	0,48932	0,5297
	<i>varLineHV</i>	0,48932	0,5297
	<i>cvLineHV</i>	0,00000	0,8691
	<i>skewLineHV</i>	0,00000	0,9397
	<i>kurtLineHV</i>	0,02161	0,5776
	μ_{AC}	0,00000	0,7982
	μ_{CP}	0,00000	0,8796
	μ_{PA}	0,31682	0,6014
	<i>meanDCB</i>	0,00000	0,7494
	<i>stdDCB</i>	0,00000	0,9381
	<i>varDCB</i>	0,00000	0,9381
	<i>cvDCB</i>	0,00000	0,9637
	<i>skewDCB</i>	0,00000	0,8589
	<i>kurtDCB</i>	0,01121	0,5975
	<i>mf</i>	0,00000	0,9531

Gráfico 9 – Valores de A_z para cada atributo na análise FVC x Normal. Em azul são os atributos de Níveis de Cinza, em vermelho os atributos de textura e em verde os atributos de forma. A linha em preto representa a linha exponencial de tendência.

Tabela 28 – Valores de p-valores calculados pelo teste Wilcoxon-Mann-Whitney e os valores de A_z para cada atributo na análise FVC Maligna x FVC Benigna. Cores das condições: $p < 0,001$ e $A_z \geq 0,9$ (preenchimento cinza escuro e escrito em branco); $0,001 \leq p \leq 0,01$ e $0,8 \leq A_z \leq 0,89$ (preenchimento cinza intermediário e escrito em preto); $0,01 \leq p \leq 0,05$ e $0,7 \leq A_z \leq 0,79$ (preenchimento cinza claro e escrito em preto); $p > 0,05$ e $A_z < 0,7$ ou valor desconhecido (?) (sem preenchimento e escrito em preto).

FVC Maligna x FVC Benigna			
	Atributo	p - valor	A_z
Níveis de Cinza	<i>cv</i>	0,30232	0,5803
	<i>skew</i>	0,00000	0,8613
	<i>kurt</i>	0,00000	0,8241
	<i>h₁</i>	0,00000	0,8491
Textura	<i>h₂</i>	0,00000	0,8659
	<i>h₃</i>	0,92529	0,4801
	<i>h₄</i>	0,00000	0,8741
	<i>h₅</i>	0,00000	0,8444
	<i>h₆</i>	0,00000	0,8291
	<i>h₇</i>	0,00000	0,8709
	<i>h₈</i>	0,00000	0,8537
	<i>h₉</i>	0,00000	0,8584
	<i>h₁₀</i>	0,00000	0,8712
	<i>h₁₁</i>	0,00000	0,8683
	<i>h₁₂</i>	0,00021	0,7307
	<i>h₁₃</i>	0,00000	0,8539
	<i>h₁₄</i>	0,18926	0,5658
	<i>C₀</i>	0,00040	0,722
Forma	<i>FDF</i>	0,00000	0,8365
	<i>CD</i>	0,00087	0,6999
	<i>M₁</i>	0,21284	0,5671
	<i>M₂</i>	0,64397	0,5176
	<i>M₃</i>	0,01092	0,6553
	<i>M₄</i>	0,04484	0,6168
	<i>M₅</i>	0,83246	0,3885
	<i>M₆</i>	0,68957	0,4797
	<i>M₇</i>	0,47299	0,5376
	<i>meanLineV</i>	0,00038	0,7026
	<i>stdLineV</i>	0,65361	0,5314
	<i>varLineV</i>	0,65361	0,5314
	<i>cvLineV</i>	0,00029	0,7208
	<i>skewLineV</i>	0,06951	0,6162
	<i>kurtLineV</i>	0,19152	0,572
	<i>meanLineH</i>	0,00000	0,7838
	<i>stdLineH</i>	0,00164	0,6774
	<i>varLineH</i>	0,00164	0,6774
	<i>cvLineH</i>	0,02572	0,6438
	<i>skewLineH</i>	0,00882	0,6379
	<i>kurtLineH</i>	0,32483	0,5458
	<i>meanLineHV</i>	0,00003	0,735
	<i>stdLineHV</i>	0,00015	0,7186
	<i>varLineHV</i>	0,00015	0,7186
	<i>cvLineHV</i>	0,09942	0,597
	<i>skewLineHV</i>	0,46945	0,5404
	<i>kurtLineHV</i>	0,77847	0,4845
	μAC	0,18926	0,5662
	μCP	0,36588	0,5646
	μPA	0,01051	0,6607
	<i>meanDCB</i>	0,00047	0,6971
	<i>stdDCB</i>	0,45316	0,5269
	<i>varDCB</i>	0,45316	0,5269
	<i>cvDCB</i>	0,36945	0,5584
	<i>skewDCB</i>	0,24657	0,573
	<i>kurtDCB</i>	0,16766	0,569
	<i>mf</i>	0,07265	0,5977

Gráfico 10 – Valores de A_z para cada atributo na análise FVC Maligna x FVC Benigna. Em azul são os atributos de Níveis de Cinza, em vermelho os atributos de textura e em verde os atributos de forma. A linha em preto representa a linha exponencial de tendência.

6.2. Seleções de atributos

i. Atributos com $p < 0,001$ e cinco maiores valores de A_z ($5+ A_z$)

As Tabelas 6 e 7 trazem os valores de A_z dos cinco atributos com maiores A_z com $p < 0,001$ que foram selecionados para compor o vetor de características dos corpos vertebrais nas classificações de “FVC x Normal” e “FVC Maligna x FVC Benigna” respectivamente. Os Gráficos 3 e 4 ilustram as curvas ROC desses cinco atributos selecionados para as classificações de “FVC x Normal” e “FVC Maligna x FVC Benigna” respectivamente.

Tabela 29 – Cinco atributos com maior valor de A_z selecionados para compor o vetor de características dos corpos vertebrais na classificação de FVC x Normal.

FVC x Normal	
Atributos	A_z
<i>cvDCB</i>	0,9637
M_1	0,9636
<i>mf</i>	0,9531
<i>skewLineHV</i>	0,9397
<i>stdDCB</i>	0,9381

Gráfico 11 – Ilustração das curvas ROC dos cinco atributos apresentados na Tabela 6 para a classificação de FVC x Normal.

Tabela 30 – Cinco atributos com maior valor de A_z selecionados para compor o vetor de características dos corpos vertebrais na classificação de FVC Maligna x FVC Benigna.

FVC Maligna x FVC Benigna	
Atributos	A_z
h_4	0,8741
h_{10}	0,8712
h_7	0,8709
h_{11}	0,8683
h_2	0,8659

Gráfico 12 – Ilustração das curvas ROC dos cinco atributos apresentados na Tabela 7 para a classificação de FVC Maligna x FVC Benigna.

ii. Atributos com $p < 0,001$ e dez maiores valores de A_z ($10+ A_z$)

As Tabelas 8 e 9 trazem os valores de A_z dos dez atributos com maiores A_z com $p < 0,001$ que foram selecionados para compor o vetor de características dos corpos vertebrais nas classificações de “FVC x Normal” e “FVC Maligna x FVC Benigna” respectivamente.

Tabela 31 – Dez atributos com maior valor de A_z selecionados para compor o vetor de características dos corpos vertebrais na classificação de FVC x Normal.

FVC x Normal	
Atributos	A_z
<i>cvDCB</i>	0,9637
M_1	0,9636
<i>mf</i>	0,9531
<i>skewLineHV</i>	0,9397
<i>stdDCB</i>	0,9381
<i>varDCB</i>	0,9381
M_4	0,936
M_2	0,9315
<i>meanLineV</i>	0,9019
<i>skewLineH</i>	0,894

Tabela 32 – Dez atributos com maior valor de A_z selecionados para compor o vetor de características dos corpos vertebrais na classificação de FVC Maligna x FVC Benigna.

FVC Maligna x FVC Benigna	
Atributos	A_z
h_4	0,8741
h_{10}	0,8712
h_7	0,8709
h_{11}	0,8683
h_2	0,8659
<i>skew</i>	0,8613
h_9	0,8584
h_{13}	0,8539
h_8	0,8537
h_1	0,8491

iii. Método automático *Wrapper*

As Tabelas 10, 11 e 12 trazem os resultados das seleções de atributos pelo método *Wrapper*, com *cross-validation* de *folds* = 10, método de busca *best first* para cada classificador e para as análises “FVC x Normal”, “FVC Maligna x FVC Benigna” e “FVC Maligna x FVC Benigna x Normal” respectivamente. Os atributos que foram selecionados 30% dos *folds* ou mais são os atributos que irão compor os vetores características do respectivo classificador.

Tabela 33 – Seleções de atributos pelo método *Wrapper*, com *cross-validation* de *folds* = 10, método de busca *best first* para cada classificador e para a análise “FVC x Normal”. Os atributos selecionados 30% dos *folds* ou mais são os atributos que irão compor os vetores características do respectivo classificador e estão destacados em negrito. Em (?) são valores não obtidos.

Seleção de atributos (<i>Wrapper</i>) - FVC x Normal								
	<i>k</i> -NN					<i>Naïve Bayes</i>	<i>RBF network</i>	<i>SVM</i>
	J48	<i>k</i> = 5	<i>k</i> = 7	<i>k</i> = 9	<i>k</i> = 11			
<i>cv</i>	0	20	20	30	20	40	80	?
<i>skew</i>	20	20	0	0	0	20	10	?
<i>kurt</i>	0	10	20	0	10	0	0	?
<i>h</i> ₁	20	20	30	40	30	0	30	?
<i>h</i> ₂	0	0	10	0	10	40	10	?
<i>h</i> ₃	50	20	10	20	10	30	50	?
<i>h</i> ₄	20	0	10	0	10	20	50	?
<i>h</i> ₅	10	0	0	0	0	30	30	?
<i>h</i> ₆	50	0	0	0	0	100	40	?
<i>h</i> ₇	10	0	0	10	0	0	20	?
<i>h</i> ₈	20	10	0	0	0	20	20	?
<i>h</i> ₉	0	30	20	10	10	50	10	?
<i>h</i> ₁₀	10	0	0	0	0	50	0	?
<i>h</i> ₁₁	0	0	10	0	10	20	40	?
<i>h</i> ₁₂	0	0	0	0	0	0	0	?
<i>h</i> ₁₃	0	0	10	0	10	0	0	?
<i>h</i> ₁₄	10	10	0	0	0	10	10	?
<i>C</i> ₀	20	0	0	0	0	0	0	?
<i>FDF</i>	10	20	10	0	10	40	50	?
<i>CD</i>	20	0	0	10	0	0	0	?
<i>M</i> ₁	30	40	70	70	70	0	40	?
<i>M</i> ₂	60	10	10	0	10	0	0	?
<i>M</i> ₃	20	10	40	20	20	0	30	?
<i>M</i> ₄	0	50	50	30	40	0	40	?
<i>M</i> ₅	0	0	0	0	0	0	30	?
<i>M</i> ₆	10	20	10	10	10	0	20	?
<i>M</i> ₇	0	0	0	0	0	0	20	?
<i>meanLineV</i>	0	0	0	0	0	0	0	?
<i>stdLineV</i>	10	10	0	0	0	10	10	?
<i>varLineV</i>	0	0	0	0	0	0	20	?
<i>cvLineV</i>	10	0	0	0	0	40	0	?
<i>skewLineV</i>	10	0	0	0	0	0	0	?
<i>kurtLineV</i>	0	0	0	0	0	0	10	?
<i>meanLineH</i>	0	0	0	0	0	30	10	?
<i>stdLineH</i>	30	0	0	0	0	20	60	?
<i>varLineH</i>	0	0	0	0	0	0	0	?
<i>cvLineH</i>	0	0	0	0	0	10	30	?
<i>skewLineH</i>	10	30	10	0	0	10	50	?
<i>kurtLineH</i>	10	10	10	0	10	40	50	?
<i>meanLineHV</i>	0	0	0	0	0	0	0	?
<i>stdLineHV</i>	0	0	0	0	0	50	0	?
<i>varLineHV</i>	0	0	0	0	0	10	10	?
<i>cvLineHV</i>	10	0	0	0	0	0	0	?
<i>skewLineHV</i>	20	0	0	10	0	20	20	?
<i>kurtLineHV</i>	0	0	0	0	0	0	40	?
μ AC	20	0	0	0	0	0	0	?
μ CP	0	10	0	0	0	10	50	?
μ PA	0	0	0	0	0	30	10	?
<i>meanDCB</i>	10	0	0	0	0	0	10	?
<i>stdDCB</i>	0	0	0	0	0	0	0	?
<i>varDCB</i>	0	0	0	0	0	0	0	?
<i>cvDCB</i>	20	40	20	10	10	100	60	?
<i>skewDCB</i>	20	0	0	0	0	50	40	?
<i>kurtDCB</i>	0	20	20	0	10	0	0	?
<i>mf</i>	10	40	10	20	10	0	10	?

Tabela 34 – Seleções de atributos pelo método *Wrapper*, com *cross-validation* de *folds* = 10, método de busca *best first* para cada classificador e para a análise “FVC Maligna x FVC Benigna”. Os atributos selecionados 30% dos *folds* ou mais são os que irão compor os vetores características do respectivo classificador e estão destacados em negrito. Em (?) são valores não obtidos.

Seleção de atributos (<i>Wrapper</i>) - FVC Maligna x FVC Benigna								
	<i>k</i> -NN					Naïve Bayes	RBF network	SVM
	J48	<i>k</i> = 5	<i>k</i> = 7	<i>k</i> = 9	<i>k</i> = 11			
<i>cv</i>	10	0	10	20	10	10	30	?
<i>skew</i>	20	0	0	0	0	50	40	?
<i>kurt</i>	0	0	10	10	10	0	0	?
<i>h</i> ₁	10	50	10	20	30	0	0	?
<i>h</i> ₂	0	40	30	20	20	10	20	?
<i>h</i> ₃	30	70	30	40	20	60	70	?
<i>h</i> ₄	0	0	0	0	0	0	10	?
<i>h</i> ₅	0	10	0	0	10	20	10	?
<i>h</i> ₆	50	30	30	20	20	40	50	?
<i>h</i> ₇	0	0	0	0	10	0	0	?
<i>h</i> ₈	10	10	10	0	0	0	0	?
<i>h</i> ₉	20	0	10	0	20	0	0	?
<i>h</i> ₁₀	0	50	10	60	50	10	20	?
<i>h</i> ₁₁	40	10	0	0	10	10	0	?
<i>h</i> ₁₂	10	0	0	0	0	0	10	?
<i>h</i> ₁₃	10	30	50	60	50	50	10	?
<i>h</i> ₁₄	0	30	20	20	20	20	80	?
<i>C</i> ₀	20	0	0	0	0	20	10	?
<i>FDF</i>	50	70	90	100	90	60	80	?
<i>CD</i>	0	0	0	0	0	50	40	?
<i>M</i> ₁	10	0	0	0	0	0	0	?
<i>M</i> ₂	0	10	10	0	0	10	20	?
<i>M</i> ₃	0	20	0	0	0	0	0	?
<i>M</i> ₄	0	40	0	30	0	0	0	?
<i>M</i> ₅	10	20	20	30	30	0	0	?
<i>M</i> ₆	0	70	20	20	10	0	0	?
<i>M</i> ₇	30	20	10	30	20	0	0	?
<i>meanLineV</i>	10	0	0	0	0	0	10	?
<i>stdLineV</i>	0	0	0	0	0	20	20	?
<i>varLineV</i>	0	0	10	0	0	0	0	?
<i>cvLineV</i>	0	10	20	10	10	30	20	?
<i>skewLineV</i>	0	0	0	0	0	10	10	?
<i>kurtLineV</i>	10	10	10	0	10	0	20	?
<i>meanLineH</i>	0	10	0	0	0	30	10	?
<i>stdLineH</i>	0	0	0	0	0	0	10	?
<i>varLineH</i>	0	0	0	0	0	0	0	?
<i>cvLineH</i>	20	0	0	10	0	10	10	?
<i>skewLineH</i>	0	0	0	0	0	10	0	?
<i>kurtLineH</i>	10	0	0	0	0	0	10	?
<i>meanLineHV</i>	0	0	0	0	0	10	10	?
<i>stdLineHV</i>	0	0	0	0	0	10	10	?
<i>varLineHV</i>	0	0	0	0	0	0	0	?
<i>cvLineHV</i>	0	0	0	10	0	10	10	?
<i>skewLineHV</i>	0	0	0	0	0	0	0	?
<i>kurtLineHV</i>	0	0	0	0	0	10	0	?
μ AC	20	0	0	0	0	20	0	?
μ CP	0	0	0	0	0	10	10	?
μ PA	0	0	0	0	0	20	10	?
<i>meanDCB</i>	0	0	0	0	0	10	40	?
<i>stdDCB</i>	0	0	0	0	0	0	10	?
<i>varDCB</i>	0	10	20	0	0	10	20	?
<i>cvDCB</i>	10	0	0	0	0	10	0	?
<i>skewDCB</i>	0	0	0	0	0	20	0	?
<i>kurtDCB</i>	10	0	0	0	0	0	0	?
<i>mf</i>	0	0	0	0	0	0	0	?

Tabela 35 – Seleções de atributos pelo método *Wrapper*, com *cross-validation* de *folds* = 10, método de busca *best first* para cada classificador e para a análise “FVC Maligna x FVC Benigna x Normal”. Os atributos selecionados 30% dos *folds* ou mais são os que irão compor os vetores características do respectivo classificador e estão destacados em negrito. Em (?) são valores não obtidos.

Seleção de atributos (<i>Wrapper</i>) - FVC Maligna x FVC Benigna x Normal								
	k-NN							
	J48	k = 5	k = 7	k = 9	k = 11	Naïve Bayes	RBF network	SVM
cv	?	0	20	10	20	10	20	?
skew	?	0	0	0	0	10	10	?
kurt	?	20	20	20	10	0	0	?
h ₁	?	0	30	10	0	0	10	?
h ₂	?	40	20	20	20	10	20	?
h ₃	?	30	30	30	10	60	70	?
h ₄	?	20	30	20	30	0	0	?
h ₅	?	10	10	0	0	30	40	?
h ₆	?	20	10	20	40	80	40	?
h ₇	?	50	20	20	50	0	0	?
h ₈	?	20	30	20	30	90	40	?
h ₉	?	0	0	0	10	10	0	?
h ₁₀	?	10	10	0	10	0	10	?
h ₁₁	?	30	20	20	20	30	30	?
h ₁₂	?	10	0	0	10	40	20	?
h ₁₃	?	30	20	20	20	20	10	?
h ₁₄	?	30	10	10	10	20	50	?
C ₀	?	0	10	10	0	0	30	?
FDF	?	20	30	30	10	70	40	?
CD	?	10	10	10	0	10	0	?
M ₁	?	20	10	30	10	20	30	?
M ₂	?	10	10	30	10	0	20	?
M ₃	?	30	20	20	30	10	30	?
M ₄	?	30	20	20	40	0	60	?
M ₅	?	0	40	40	10	0	10	?
M ₆	?	50	20	30	20	0	0	?
M ₇	?	10	0	10	10	10	10	?
meanLineV	?	20	20	20	20	30	20	?
stdLineV	?	0	0	10	10	10	10	?
varLineV	?	10	10	0	0	10	20	?
cvLineV	?	10	0	10	10	60	20	?
skewLineV	?	0	10	0	0	10	0	?
kurtLineV	?	0	0	0	0	0	10	?
meanLineH	?	0	0	0	0	60	30	?
stdLineH	?	0	0	0	0	30	0	?
varLineH	?	0	0	0	0	0	20	?
cvLineH	?	0	0	0	10	20	20	?
skewLineH	?	20	10	0	0	20	40	?
kurtLineH	?	10	0	10	10	10	20	?
meanLineHV	?	10	10	30	20	10	50	?
stdLineHV	?	0	0	0	0	40	0	?
varLineHV	?	0	0	0	0	0	20	?
cvLineHV	?	10	30	30	40	0	0	?
skewLineHV	?	10	20	10	20	50	10	?
kurtLineHV	?	10	30	30	30	10	20	?
μAC	?	0	0	0	0	30	30	?
μCP	?	20	0	0	10	60	40	?
μPA	?	0	0	0	0	40	10	?
meanDCB	?	0	0	0	0	20	20	?
stdDCB	?	50	20	40	30	20	30	?
varDCB	?	40	50	40	30	30	50	?
cvDCB	?	20	30	30	30	0	0	?
skewDCB	?	10	20	0	0	60	50	?
kurtDCB	?	0	0	0	0	0	0	?
mf	?	0	0	0	0	70	20	?

6.3. Classificação

6.3.1. Classificação segundo os métodos de seleção de atributos

A seguir estão os resultados obtidos nas classificações para cada método de seleção de atributos.

i. Todos os atributos

As Tabelas 13, 14 e 15 a seguir trazem os resultados das classificações para as análises “FVC x Normal”, “FVC Maligna x FVC Benigna” e “FVC Maligna x FVC Benigna x Normal” respectivamente utilizando o vetor de características com todos os atributos calculados para os corpos vertebrais.

Tabela 36 – Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com *cross-validation* de 10-folds, utilizando o vetor de características com todos os atributos calculados para os corpos vertebrais para a análise “FVC x Normal”. Em negrito estão os valores das médias aritméticas dos resultados das classificações.

FVC x Normal - Todos os atributos					
	Classificador	<i>acc</i>	AUROC	<i>sens</i>	<i>spec</i>
k-NN	J48	0,864	0,842	0,863	0,865
	<i>k</i> = 5	0,916	0,954	0,892	0,944
	<i>k</i> = 7	0,911	0,955	0,882	0,944
	<i>k</i> = 9	0,911	0,957	0,882	0,944
	<i>k</i> = 11	0,911	0,957	0,882	0,944
	Naïve Bayes	0,911	0,973	0,853	0,978
	RBF network	0,911	0,92	0,863	0,966
	SVM	0,895	0,896	0,892	0,899
	Média dos resultados	0,904	0,932	0,876	0,935

Tabela 37 – Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com *cross-validation* de 10-folds, utilizando o vetor de características com todos os atributos calculados para os corpos vertebrais para a análise “FVC Maligna x FVC Benigna”. Em negrito estão os valores das médias aritméticas dos resultados das classificações.

FVC Maligna x FVC Benigna - Todos os atributos					
	Classificador	<i>acc</i>	AUROC	<i>sens</i>	<i>spec</i>
k-NN	J48	0,794	0,687	0,735	0,849
	<i>k</i> = 5	0,716	0,797	0,694	0,736
	<i>k</i> = 7	0,735	0,821	0,755	0,717
	<i>k</i> = 9	0,735	0,816	0,735	0,736
	<i>k</i> = 11	0,745	0,855	0,755	0,736
	Naïve Bayes	0,765	0,879	0,878	0,66
	RBF network	0,804	0,778	0,918	0,698
	SVM	0,696	0,698	0,735	0,66
	Média dos resultados	0,749	0,791	0,776	0,724

Tabela 38 – Acurácia, AUROC de cada classificação, com *cross-validation* de 10-folds, utilizando o vetor de características com todos os atributos calculados para os corpos vertebrais para a análise “FVC Maligna x FVC Benigna x Normal”. Em negrito estão os valores das médias aritméticas dos resultados das classificações.

FVC Maligna x FVC Benigna x Normal Todos os atributos			
	Classificador	<i>acc</i>	AUROC
<i>k</i> -NN	J48	0,812	0,833
	<i>k</i> = 5	0,806	0,913
	<i>k</i> = 7	0,822	0,925
	<i>k</i> = 9	0,817	0,928
	<i>k</i> = 11	0,817	0,928
	<i>Naïve Bayes</i>	0,812	0,929
	<i>RBF network</i>	0,827	0,902
	SVM	0,822	0,864
	Média dos resultados	0,817	0,903

ii. Atributos com $p < 0,001$ e cinco maiores valores de A_z ($5+ A_z$)

As Tabelas 16 e 17 a seguir trazem os resultados das classificações para as análises “FVC x Normal” e “FVC Maligna x FVC Benigna” respectivamente utilizando o vetor de características com os cinco atributos de maiores valores de A_z .

Tabela 39 – Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com *cross-validation* de 10-folds, utilizando o vetor de características com os cinco atributos de maiores valores de A_z calculados para a análise “FVC x Normal”. Em negrito estão os valores das médias aritméticas dos resultados das classificações.

FVC x Normal – 5+ A_z					
	Classificador	<i>acc</i>	AUROC	<i>sens</i>	<i>spec</i>
<i>k</i> -NN	J48	0,859	0,886	0,814	0,91
	<i>k</i> = 5	0,874	0,949	0,853	0,899
	<i>k</i> = 7	0,869	0,946	0,853	0,888
	<i>k</i> = 9	0,869	0,95	0,853	0,888
	<i>k</i> = 11	0,879	0,949	0,853	0,91
	<i>Naïve Bayes</i>	0,901	0,959	0,863	0,944
	<i>RBF network</i>	0,895	0,931	0,853	0,944
	SVM	0,895	0,897	0,921	0,873
	Média dos resultados	0,880	0,933	0,858	0,907

Tabela 40 – Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com *cross-validation* de 10-folds, utilizando o vetor de características com os cinco atributos de maiores valores de A_z calculados para a análise “FVC Maligna x FVC Benigna”. Em negrito estão os valores das médias aritméticas dos resultados das classificações.

FVC Maligna x FVC Benigna – 5+ A_z					
	Classificador	acc	AUROC	sens	spec
k-NN	J48	0,667	0,729	0,714	0,623
	$k = 5$	0,775	0,809	0,735	0,811
	$k = 7$	0,775	0,809	0,735	0,811
	$k = 9$	0,775	0,809	0,735	0,811
	$k = 11$	0,775	0,809	0,735	0,811
	Naïve Bayes	0,735	0,818	0,918	0,566
	RBF network	0,735	0,795	0,735	0,736
	SVM	0,588	0,584	0,698	0,469
	Média dos resultados	0,728	0,770	0,751	0,705

iii. Atributos com $p < 0,001$ e dez maiores valores de A_z (10+ A_z)

As Tabelas 18 e 19 a seguir trazem os resultados das classificações para as análises “FVC x Normal” e “FVC Maligna x FVC Benigna” respectivamente utilizando o vetor de características com os dez atributos de maiores valores de A_z .

Tabela 41 – Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com *cross-validation* de 10-folds, utilizando o vetor de características com os dez atributos de maiores valores de A_z calculados para a análise “FVC x Normal”. Em negrito estão os valores das médias aritméticas dos resultados das classificações.

FVC x Normal – 10+ A_z					
	Classificador	acc	AUROC	sens	spec
k-NN	J48	0,869	0,885	0,843	0,899
	$k = 5$	0,874	0,931	0,863	0,888
	$k = 7$	0,895	0,939	0,873	0,921
	$k = 9$	0,901	0,943	0,892	0,91
	$k = 11$	0,901	0,942	0,902	0,899
	Naïve Bayes	0,879	0,96	0,824	0,944
	RBF network	0,885	0,936	0,824	0,955
	SVM	0,885	0,887	0,921	0,853
	Média dos resultados	0,886	0,928	0,868	0,909

Tabela 42 – Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com cross-validation de 10-folds, utilizando o vetor de características com os dez atributos de maiores valores de A_z calculados para a análise “FVC Maligna x FVC Benigna”. Em negrito estão os valores das médias aritméticas dos resultados das classificações.

FVC Maligna x FVC Benigna – 10+ A_z					
	Classificador	acc	AUROC	sens	spec
k-NN	J48	0,637	0,666	0,592	0,679
	$k = 5$	0,696	0,751	0,633	0,755
	$k = 7$	0,696	0,751	0,633	0,755
	$k = 9$	0,696	0,751	0,633	0,755
	$k = 11$	0,696	0,751	0,633	0,755
	Naïve Bayes	0,755	0,846	0,878	0,642
	RBF network	0,716	0,797	0,776	0,66
	SVM	0,676	0,675	0,717	0,633
	Média dos resultados	0,696	0,749	0,687	0,704

iv. Método automático *Wrapper*

As Tabelas 20, 21 e 22 a seguir trazem os resultados das classificações para as análises “FVC x Normal”, “FVC Maligna x FVC Benigna” e “FVC Maligna x FVC Benigna x Normal” respectivamente utilizando o vetor de características específico de cada classificador obtido na seleção de atributos com o método *Wrapper*.

Tabela 43 – Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com cross-validation de 10-folds, utilizando o vetor de características específico de cada classificador, obtido na seleção de atributo com o método *Wrapper*, calculados para a análise “FVC x Normal”. Em negrito estão os valores das médias aritméticas dos resultados das classificações e os valores em (?) são valores não obtidos.

FVC x Normal – <i>Wrapper</i>					
	Classificador	acc	AUROC	sens	spec
k-NN	J48	0,916	0,926	0,922	0,91
	$k = 5$	0,906	0,941	0,882	0,933
	$k = 7$	0,916	0,948	0,912	0,921
	$k = 9$	0,901	0,942	0,882	0,921
	$k = 11$	0,906	0,954	0,912	0,899
	Naïve Bayes	0,958	0,984	0,941	0,978
	RBF network	0,932	0,954	0,892	0,978
	SVM	?	?	?	?
	Média dos resultados	0,919	0,950	0,906	0,934

Tabela 44 – Acurácia, AUROC, sensibilidade e especificidade de cada classificação, com cross-validation de 10-folds, utilizando o vetor de características específico de cada classificador, obtido na seleção de atributo com o método *Wrapper*, calculados para a análise “FVC Maligna x FVC Benigna”. Em negrito estão os valores das médias aritméticas dos resultados das classificações e os valores em (?) são valores não obtidos.

FVC Maligna x FVC Benigna – <i>Wrapper</i>					
	Classificador	<i>acc</i>	AUROC	<i>sens</i>	<i>spec</i>
<i>k</i> -NN	J48	0,892	0,879	0,918	0,868
	<i>k</i> = 5	0,794	0,867	0,776	0,811
	<i>k</i> = 7	0,824	0,836	0,816	0,83
	<i>k</i> = 9	0,814	0,899	0,857	0,774
	<i>k</i> = 11	0,824	0,88	0,837	0,811
	<i>Naïve Bayes</i>	0,863	0,93	0,837	0,887
	<i>RBF network</i>	0,824	0,856	0,776	0,868
	SVM	?	?	?	?
	Média dos resultados	0,834	0,878	0,831	0,836

Tabela 45 – Acurácia e AUROC de cada classificação, com cross-validation de 10-folds, utilizando o vetor de características específico de cada classificador, obtido na seleção de atributo com o método *Wrapper*, calculados para a análise “FVC Maligna x FVC Benigna x Normal”. Em negrito estão os valores das médias aritméticas dos resultados das classificações e os valores em (?) são valores não obtidos.

FVC Maligna x FVC Benigna x Normal – <i>Wrapper</i>			
	Classificador	<i>acc</i>	AUROC
<i>k</i> -NN	J48	?	?
	<i>k</i> = 5	0,801	0,879
	<i>k</i> = 7	0,843	0,946
	<i>k</i> = 9	0,785	0,899
	<i>k</i> = 11	0,806	0,932
	<i>Naïve Bayes</i>	0,843	0,953
	<i>RBF network</i>	0,827	0,931
	SVM	?	?
	Média dos resultados	0,818	0,923

6.3.2. Classificações dos radiologistas e do computador

Como o método *Wrapper* de seleção de atributos obteve a maior média dos valores de AUROC, os resultados das classificações utilizando esse método de seleção de atributos foram comparados com classificações realizadas por três radiologistas. As Tabelas 23 e 24 trazem os valores de sensibilidade e especificidade, com seus respectivos intervalos de confiança de 95%, para as classificações feitas pelos radiologistas e pelos classificadores para as análises “FVC x Normal” e “FVC Maligna x FVC Benigna” respectivamente.

Tabela 46 – Resultados de sensibilidade e especificidade, com seus respectivos intervalos de confiança de 95%, para as classificações feitas pelos radiologistas e pelos classificadores para as análises “FVC x Normal”.

FVC x Normal – Radiologistas e computador				
		Valor	Intervalo de confiança de 95%	
			Limite inferior	Limite Superior
Radiologista 1	<i>sens</i>	0,882	0,800	0,935
	<i>spec</i>	0,888	0,799	0,942
Radiologista 2	<i>sens</i>	0,863	0,777	0,920
	<i>spec</i>	0,978	0,914	0,996
Radiologista 3	<i>sens</i>	0,990	0,939	0,999
	<i>spec</i>	0,472	0,366	0,580
J48	<i>sens</i>	0,925	0,852	0,964
	<i>spec</i>	0,910	0,826	0,958
<i>k</i> = 5	<i>sens</i>	0,882	0,800	0,935
	<i>spec</i>	0,933	0,854	0,972
<i>k</i> = 7	<i>sens</i>	0,912	0,835	0,956
	<i>spec</i>	0,921	0,839	0,965
<i>k</i> = 9	<i>sens</i>	0,882	0,800	0,935
	<i>spec</i>	0,921	0,839	0,965
<i>k</i> = 11	<i>sens</i>	0,912	0,835	0,956
	<i>spec</i>	0,899	0,812	0,950
Naïve Bayes	<i>sens</i>	0,941	0,871	0,976
	<i>spec</i>	0,978	0,914	0,996
RBF network	<i>sens</i>	0,892	0,811	0,942
	<i>spec</i>	0,978	0,914	0,996

Gráfico 13 – Ilustração dos resultados de sensibilidade, com intervalos de confiança de 95%, para “FVC x Normal” das avaliações dos três radiologistas e dos classificadores com método automático Wrapper de seleção de atributos de acordo com a Tabela 23.

Gráfico 14 – Ilustração dos resultados de especificidade, com intervalos de confiança de 95%, para “FVC x Normal” das avaliações dos três radiologistas e dos classificadores com método automático Wrapper de seleção de atributos de acordo com a Tabela 23.

Tabela 47 – Resultados de sensibilidade e especificidade, com seus respectivos intervalos de confiança de 95%, para as classificações feitas pelos radiologistas e pelos classificadores para as análises “FVC Maligna x FVC Benigna”.

FVC Maligna x FVC Benigna – Radiologistas e computador					
		Valor	Intervalo de confiança de 95%		
			Limite inferior	Limite Superior	
k - NN	Radiologista 1	sens	0,886	0,746	0,957
		spec	0,804	0,656	0,901
	Radiologista 2	sens	0,675	0,508	0,809
		spec	0,729	0,579	0,843
	Radiologista 3	sens	0,714	0,565	0,830
		spec	0,923	0,806	0,975
	J48	sens	0,918	0,795	0,974
		spec	0,868	0,740	0,941
	k = 5	sens	0,776	0,630	0,878
		spec	0,811	0,676	0,901
	k = 7	sens	0,816	0,675	0,908
		spec	0,830	0,697	0,915
	k = 9	sens	0,857	0,721	0,936
		spec	0,774	0,635	0,873
	k = 11	sens	0,837	0,698	0,922
		spec	0,811	0,676	0,901
	Naïve Bayes	sens	0,837	0,698	0,922
		spec	0,887	0,763	0,953
	RBF network	sens	0,776	0,630	0,878
		spec	0,868	0,740	0,941

Gráfico 15 – Ilustração dos resultados de sensibilidade, com intervalos de confiança de 95%, para “FVC Maligna x FVC Benigna” das avaliações dos três radiologistas e dos classificadores com método automático Wrapper de seleção de atributos de acordo com a Tabela 24.

Gráfico 16 – Ilustração dos resultados de especificidade, com intervalos de confiança de 95%, para “FVC Maligna x FVC Benigna” das avaliações dos três radiologistas e dos classificadores com método automático Wrapper de seleção de atributos de acordo com a Tabela 24.

7. DISCUSSÃO

7.1. Seleção de atributos

i. Seleção manual dos atributos

Genant et al. (1993) classificaram as fraturas de corpos vertebrais utilizando as medidas de alturas anteriores, centrais e posteriores dos corpos vertebrais. Diferenças entre as alturas acima de 20% já são consideradas como FVCs. Quando os valores das alturas são muito menores que os valores da largura do corpo vertebral, a possibilidade de ter ocorrido uma FVC é real. Informações sobre abaulamentos e convexidades posteriores laterais indicam suspeitas de malignidade. Por este motivo desenvolvemos atributos de forma que trabalham com essa relação de altura e largura do corpo vertebral como os atributos estatísticos de altura e largura vertebral utilizando o eixo principal e os atributos estatísticos da distância do centroide vertebral à sua borda.

No caso de FVCs benignas, os platôs vertebrais podem adquirir um formato côncavo mais acentuado e a impactação do osso subcondral pode resultar em contornos ásperos com reentrâncias. FVCs benignas podem causar retropulsão de fragmentos da parede posterior com contornos angulados ou irregulares (CUENOD *et al.*, 1996; JUNG *et al.*, 2003). Este aspecto nos motivou a programar atributos de forma como C_o , CD e FDF que analisam os contornos vertebrais e suas concavidades. Já os momentos invariantes de Hu (HU, 1962) são atributos que auxiliam em classificações de objetos com vários tipos de formas e por isso foi incluído no trabalho.

A intensidade de sinal no corpo vertebral é um importante critério para discriminar FVCs benignas e malignas em IRM, por isso foram extraídos os atributos estatísticos de níveis de cinza calculados a partir do histograma da ROI. A fim de analisar a distribuição dos níveis de cinza, os clássicos 14 descritores de Haralick et al. (1973) foram aplicados ao trabalho.

Esses critérios nos levaram ao vetor de característica de 55 atributos extraídos das ROIs.

ii. Seleção utilizando p-valor e A_z

Ao analisar a Tabela 4, observa-se que 31 dos 38 atributos de forma obtiveram um $p < 0,05$ e desses 31 atributos, 27 obtiveram um $p < 0,001$ para análise de FVC x Normal. Para a mesma análise, 12 atributos dos 14 de textura obtiveram o $p < 0,05$ sendo três com $p < 0,001$. Além disso, o atributo de nível de cinza *cv* obteve $p < 0,001$.

O atributo de nível de cinza *cv* obteve uma $A_z = 0,7507$, enquanto o maior valor de A_z dos atributos de textura foi obtido pelo *h6* com $A_z = 0,73$. Os demais atributos de textura com $p < 0,05$, porém, obtiveram valor de $A_z < 0,68$. Já os atributos de forma obtiveram elevados valores de A_z comparados aos atributos de textura e níveis de cinza. Nove atributos de forma obtiveram $A_z > 0,90$, sendo o atributo *cvDCB* o que obteve maior valor ($A_z = 0,9637$). Não foi possível obter os valores de A_z para os atributos M_5 e M_7 devido a erros ocorridos no software ROCKIT.

O Gráfico 1 ilustra os valores de A_z de todos os atributos para a análise FVC x Normal. Nota-se o crescimento da curva exponencial de tendência ao se aproximar dos valores de A_z do conjunto de atributos de forma.

Esses resultados mostram que os atributos de forma desenvolvidos nesse trabalho possuem grande potencial na detecção de FVCs, porém, apesar de obterem valores de A_z inferiores, atributos de textura e níveis de cinza obtiveram alta significatividade nos testes estatísticos. Isso indica que esses atributos também tem potencial na detecção de FVCs.

Na Tabela 5, observa-se que 12 dos 14 atributos de textura e os atributos de níveis de cinza *skew* e *kurt* obtiveram um $p < 0,001$ para a análise de FVC Maligna x FVC Benigna. Quanto aos atributos de forma, 17 obtiveram $p < 0,05$, 14 a menos que para a análise anterior, e desses 17 atributos, 10 obtiveram $p < 0,001$.

O atributo de forma *FDF* obteve um valor de $A_z = 0,8365$, enquanto os demais atributos de forma com $p < 0,05$ possuem $A_z < 0,79$. Já os atributos de níveis de cinza e

textura com $p < 0,001$, com exceção de h_{12} , obtiveram $A_z > 0,82$, sendo h_4 o que obteve maior A_z ($A_z = 0,8741$).

O Gráfico 2 ilustra os valores de A_z de todos os atributos para a análise FVC Maligna x FVC Benigna. Ao contrário da análise FVC x Normal, nota-se o decaimento da curva exponencial de tendência ao se aproximar dos valores de A_z do conjunto de atributos de forma.

Esses resultados mostram que os atributos de textura e níveis de cinza usados nesse trabalho possuem grande potencial na classificação de FVCs em FVC maligna ou FVC benigna, porém, apesar de obterem valores de A_z inferiores, atributos de forma obtiveram alta significância nos testes estatísticos. Isso indica que esses atributos também tem potencial na classificação de FVC Maligna x FVC Benigna.

Vale observar que o atributo de forma FDF não seguiu a tendência dos demais atributos de forma. Ele obteve $p < 0,001$ para a análise FVC Maligna x FVC Benigna, porém $p > 0,05$ para FVC x Normal. Além disso, seu valor A_z está dentro do intervalo de valores de A_z dos atributos de textura e níveis de cinza. O mesmo pode ser notado para o atributo de níveis de cinza cv , que obteve $p < 0,001$ para FVC x Normal, porém seu $p > 0,05$ para FVC Maligna x FVC Benigna.

Essas estatísticas reforçam a hipótese de que a extração de características de forma das ROIs geraria bons atributos para a detecção de FVCs e auxiliariam também na classificação de FVCs malignas e FVCs benignas. E, Além de reforçar a outra hipótese de que a extração de características de textura e níveis de cinza das ROIs geraria bons atributos para a classificação de FVCs malignas e FVCs benignas, essa estatística gera outra hipótese de que atributos de textura e níveis de cinza também auxiliam na detecção de FVCs, ou seja, na análise FVC x Normal.

Para a análise de FVC x Normal, ao seguir o critério de seleção de atributos dos 5+ A_z com p-valor $< 0,001$, cujos valores de A_z dos cinco atributos com maiores A_z estão na Tabela 6

e a ilustração das curvas ROC estão no Gráfico 3, apenas atributos de forma compõe o seu vetor de característica. O mesmo ocorre para os 10+ Az, cujos valores de Az dos dez atributos com maiores Az estão na Tabela 8.

Para a análise FVC Maligna x FVC Benigna, ao seguir o critério de seleção de atributos dos 5+ Az com $p\text{-valor} < 0,001$, cujos valores de Az dos cinco atributos com maiores Az estão na Tabela 7 e a ilustração das curvas ROC estão no Gráfico 4, apenas atributos de textura compõe o seu vetor de característica. Para os 10+ Az, cujos valores de Az dos dez atributos com maiores Az estão na Tabela 9, seu vetor de característica foi composto por nove atributos de textura e um de níveis de cinza.

Não foram feitos os cálculos dos atributos com 5+Az e 10+Az para a análise com as três classes juntas FVC Maligna x FVC Benigna x Normal, pois o software ROCKIT realiza apenas análises binárias.

iii. Método automático *Wrapper*

A seleção de atributos pelo método *Wrapper*, realizada no software WEKA, gerou os resultados presentes nas Tabelas 10, 11 e 12 e em negrito estão os valores dos atributos que compuseram os vetores de características para cada classificador. Os valores da seleção de atributos pelo método *Wrapper* para o classificador SVM não foram obtidos devido ao tempo que tarda para realizar essa seleção. Após quatro dias da execução da seleção de atributos pelo método *Wrapper* para o classificador SVM na análise de FVC Maligna x FVC Benigna, esta ainda não havia sido concluída. Por isso foi decidido não realizar a seleção de atributos por esse método para o classificador SVM. Para a análise FVC Maligna x FVC Benigna x Normal, nota-se na Tabela 12 que os valores da seleção de atributos para o classificador J48 também não foram obtidos. Isso se deve a um erro não identificado pelo *log* do software WEKA o qual parava a execução no segundo *fold* para a seleção de atributos desse classificador.

Diferentemente dos métodos $5+A_z$ e $10+A_z$ para FVC x Normal, o método *Wrapper* selecionou atributos de textura e níveis de cinza, além de atributos de forma, para compor os vetores de características. O mesmo ocorre para a análise FVC Maligna x FVC Benigna, onde o método *Wrapper* selecionou atributos de forma, além de atributos de textura e níveis de cinza, para compor os vetores de características. Um destaque para o atributo *FDF* na Tabela 11, que apesar de ser um atributo de forma, foi selecionado pelo menos 50% para todos os classificadores para a análise FVC Maligna x FVC Benigna. Nenhum outro atributo foi selecionado por todos os classificadores.

Para a análise FVC x Normal, cada classificador selecionou em média aproximadamente oito atributos para compor seu vetor de característica dos 55 atributos calculados, sendo 20 atributos selecionados pelo classificador RBF *network* e apenas três pelo classificador *k*-NN para $k = 11$. Para a análise FVC Maligna x FVC Benigna, em média aproximadamente sete atributos foram selecionados para compor o vetor de características de cada classificador, sendo dez atributos selecionados pelo classificador *k*-NN para $k = 5$ e cinco atributos selecionados para $k = 7$, $k = 11$ e J48. Para a análise FVC Maligna x FVC Benigna x Normal, cada classificador selecionou em média aproximadamente 14 atributos para compor o vetor de características, sendo 19 atributos selecionados pelos classificadores *naïve* Bayes e RBF *network* e dez atributos selecionados pelo *k*-NN para $k = 7$.

7.2. Classificação dos algoritmos

7.2.1. Vetor de características com os 55 atributos

Os resultados das classificações feitas utilizando o vetor de característica com os 55 atributos calculados das ROIs estão nas Tabelas 13, 14 e 15.

Para a análise FVC x Normal, com resultados na Tabela 13, nota-se que os valores para acurácia são iguais entre os classificadores: *naïve* Bayes, RBF *network* e *k*-NN para $k = 7$,

9, 11. Apesar do maior valor de acurácia tenha sido de 0,916 seja para k -NN para $k = 5$ e o menor valor de acurácia de 0,864 seja para J48, a diferença é de apenas 0,052. Apesar de não ter obtido o maior valor de sensibilidade, o classificador *naïve* Bayes obteve o maior valor de $AUROC = 0,973$, área 0,131 maior que o menor valor de AUROC pertencente ao classificador J48, e obteve o maior valor de especificidade de 0,978. Os classificadores SVM e k -NN para $k = 5$ empataram com o maior valor de sensibilidade de 0,892. No geral, os classificadores obtiveram resultados aproximados e de alto desempenho para a detecção de FVCs.

Para a análise FVC Maligna x FVC Benigna, com resultados na Tabela 14, o classificador RBF *network* se destacou pelo alto valor de sensibilidade de 0,918 além do maior valor de acurácia de 0,804. Os classificadores *naïve* Bayes e k -NN para $k = 7, 9, 11$ obtiveram seus valores de $AUROC > 0,8$, sendo *naïve* Bayes novamente o classificador de maior valor de $AUROC = 0,879$. Este classificador, porém, obteve um valor baixo de especificidade de 0,66. Para essa análise, a média dos resultados não alcançou 0,8.

Para a análise com as três classes FVC Maligna x FVC Benigna x Normal, com resultados na Tabela 15, observa-se que o classificador *naïve* Bayes obteve novamente o maior valor de média ponderada de $AUROC = 0,929$ e RBF *network* o maior valor de acurácia de 0,827. Ao analisar a matriz de confusão de ambos classificadores, nota-se que ambos classificaram corretamente 97,7% dos casos normais, porém classificaram incorretamente 33% dos casos de FVC benigna como FVC maligna.

7.2.2. Atributos com $p < 0,001$ e $5+A_z$

Os resultados das classificações feitas utilizando o vetor de característica com os cinco atributos calculados das ROIs com $p < 0,001$ e $5+A_z$ estão nas Tabelas 16 e 17.

Para a análise FVC x Normal, com resultados presentes na Tabela 16, o classificador *naïve* Bayes obteve o maior valor de acurácia de 0,901, maior valor de $AUROC = 0,959$ e maior

valor de especificidade de 0,944, sendo este último valor igual ao do classificador RBF *network*. Esses valores são inferiores aos valores contidos na Tabela 13 que traz os resultados dos classificadores com o vetor de atributos com os 55 atributos. O classificador SVM obteve o maior valor de sensibilidade de 0,921, mas ao menos tempo obteve o resultado mais baixo de especificidade de 0,873. Ao comparar novamente com a Tabela 13, SVM teve um ganho de sensibilidade, porém uma perda de especificidade. Houve uma leve queda de desempenho na média dos resultados de acurácia, sensibilidade e especificidade dos classificadores em relação às classificações feitas com o vetor de características com todos os atributos (Tabela 13). Como o vetor de característica dos $5+A_z$ para FVC x Normal foi composto apenas por atributos de forma, isso fortalece a hipótese de que atributos de textura e níveis de cinza auxiliam na detecção de FVCs.

Para a análise FVC Maligna x FVC Benigna, com resultados presentes na Tabela 17, nota-se um aumento nos valores de especificidade para os classificadores *k*-NN e RBF *network* em relação aos valores da análise de FVC Maligna x FVC Benigna com todos os atributos presentes na Tabela 14. Porém os valores de especificidade diminuíram para os demais classificadores com destaque para o classificador SVM que obteve valor de especificidade menor que 0,5. Novamente o maior valor de $AUROC = 0,818$ foi para o classificador *naïve* Bayes, além do maior valor de sensibilidade de 0,918. Houve uma leve queda de desempenho na média dos resultados dos classificadores em relação às classificações feitas com o vetor de características com todos os atributos (Tabela 14). Como o vetor de característica dos $5+A_z$ para FVC Maligna x FVC Benigna foi composto apenas por atributos de textura, isso fortalece a hipótese de que atributos de forma possam auxiliar na classificação de FVCs em maligna ou benigna.

7.2.3. Atributos com $p < 0,001$ e $10+A_z$

Os resultados das classificações feitas utilizando o vetor de característica com os dez atributos calculados das ROIs com $p < 0,001$ e $10+A_z$ estão nas Tabelas 18 e 19.

Para a análise FVC x Normal, com resultados presentes na Tabela 18, o classificador *naïve* Bayes novamente obteve o maior valor de $AUROC = 0,96$, os classificadores k -NN para $k = 9$ e 11 obtiveram os maiores valores de acurácia de $0,901$, RBF *network* obteve o maior valor de especificidade de $0,955$ e novamente o classificador SVM obteve o maior valor de sensibilidade de $0,921$, porém o mais baixo dos valores de especificidade de $0,853$. A média dos resultados de classificação com o vetor de características $10+A_z$ tem valores próximos aos valores de classificação do vetor de características $5+A_z$ para a análise FVC x Normal.

Para a análise FVC Maligna x FVC Benigna, com resultados presentes na Tabela 19, o classificador *naïve* Bayes novamente obteve o maior valor de acurácia de $0,755$, $AUROC = 0,846$, sensibilidade de $0,878$ e para todos os classificadores k -NN especificidade de $0,755$. Com o vetor de características $10+A_z$ para a análise FVC Maligna x FVC Benigna, as médias dos resultados não ultrapassaram $0,75$. Assim como o vetor de característica dos $5+A_z$ para FVC Maligna x FVC Benigna que foi composto apenas por atributos de textura, $10+A_z$ não tiveram atributos de forma em sua composição. Isso fortalece a hipótese de que atributos de forma possam auxiliar na classificação de FVCs em maligna ou benigna. Afinal, a média dos resultados da análise FVC Maligna x FVC Benigna utilizando os 55 atributos (Tabela 14) foram maiores que a média dos resultados para as classificações com o vetor de características $5+A_z$ (Tabela 17) e $10+A_z$ (Tabela 19).

7.2.4. Método automático *Wrapper*

Os resultados das classificações feitas utilizando o vetor de característica com os atributos selecionados especificamente para cada classificador utilizando o método automático *Wrapper* estão nas Tabelas 20, 21 e 22.

Para a análise FVC x Normal, com resultados na Tabela 20, o classificador *naïve* Bayes obteve o melhor resultado com acurácia de 0,958, $AUROC = 0,984$, sensibilidade de 0,941 e especificidade de 0,978. Apesar da ausência dos resultados para o classificador SVM, as médias dos resultados dos classificadores obtiveram valores acima de 0,9 com valor médio de $AUROC = 0,950$. Resultados promissores para a detecção de FVCs e melhores que os apresentados anteriormente para a análise FVC x Normal.

Para a análise FVC Maligna x FVC Benigna, com resultados na Tabela 21, o classificador J48 apresentou maior valor de acurácia de 0,892 e sensibilidade de 0,918. Novamente o classificador *naïve* Bayes apresentou o maior valor de $AUROC = 0,93$ e também especificidade de 0,887. Assim como na análise anterior, apesar da ausência dos resultados para o classificador SVM, as médias dos resultados dos classificadores obtiveram valores acima de 0,83 com valor médio de $AUROC = 0,878$. Esses resultados são promissores para a classificação de FVCs em malignas ou benignas e melhores que os apresentados anteriormente para a análise FVC Maligna x FVC Benigna.

Para a análise com as três classes FVC Maligna x FVC Benigna x Normal, com resultados na Tabela 22, os classificadores *naïve* Bayes e k -NN para $k = 7$ apresentaram os maiores valores de acurácia de 0,843 e, novamente, o classificador *naïve* Bayes apresentou o maior resultado de média ponderada de $AUROC = 0,953$. Ao comparar com os resultados da Tabela 15 para o vetor de características com os 55 atributos, nota-se que o vetor de característica com atributos selecionados pelo método *Wrapper* manteve o mesmo desempenho da média de acurácia dos classificadores e aumentou em 0,02 o valor da média

das médias ponderadas de AUROC dos classificadores. De acordo com a matriz de confusão da classificação feita pelo classificador *naïve* Bayes, a quantidade de casos de FVCs benignas classificadas incorretamente como FVCs malignas caíram de 33% com vetor de características de 55 atributos para 15,1% com o método *Wrapper*.

Mesmo selecionando atributos com $p > 0,05$ para compor os vetores de características dos classificadores, o método *Wrapper* de seleção de atributos obteve as maiores médias dos resultados de AUROC dos classificadores para as análises “FVC x Normal” e “FVC Maligna x FVC Benigna”. Portanto seus resultados de acurácia, sensibilidade e especificidade foram escolhidos para serem comparados aos resultados das classificações feitas pelos três radiologistas.

7.3. Classificação dos radiologistas e computador

Os resultados dos valores estimados de sensibilidade, especificidade e seus intervalos de confiança de 95% para a análise de FVC x Normal presentes na Tabela 23 estão ilustrados respectivamente nos Gráficos 5 e 6. Já os resultados de sensibilidade, especificidade e seus intervalos de confiança de 95% para a análise de FVC Maligna x FVC Benigna presentes na Tabela 24 estão ilustrados respectivamente nos Gráficos 7 e 8.

Para a análise FVC x Normal, o Radiologista 3 obteve o maior valor de sensibilidade de 0,990 com intervalo de confiança = [0,939, 0,999]. Isso significa que o Radiologista 3 tende a detectar corretamente uma FVC em 99%. Estatisticamente o Radiologista 3 obteve maior sensibilidade do que os Radiologistas 1 e 2 e os classificadores *k*-NN para $k = 5$ e 9. Ainda sobre a sensibilidade, não houve diferença entre o Radiologista 3 e os classificadores J48, *naïve* Bayes, RBF *network* e *k*-NN para $k = 7$ e 11. Os classificadores são estatisticamente iguais quanto à sensibilidade, pois os intervalos de confiança se sobrepõem e com a boa marca de [0,8, 0,976].

Ainda sobre a análise FVC x Normal, apesar do alto valor de sensibilidade, o Radiologista 3 obteve o baixo resultado de especificidade de 0,472 com intervalo de confiança = [0,366, 0,580], enquanto os demais radiologistas e classificadores obtiveram especificidade média maior que 0,88. Mesmo tendo obtido o melhor valor de sensibilidade, seu valor de especificidade menor que 0,5 diz que a probabilidade desse radiologista classificar corretamente que um corpo vertebral está dentro da normalidade é menor que uma classificação aleatória. Os classificadores *naïve Bayes*, *RBF network* e o Radiologista 2 obtiveram o mesmo resultado de especificidade de 0,978 com intervalo de confiança = [0,914, 0,996], empatando entre os maiores valores de especificidade. Estes, portanto, tendem a detectar corretamente um corpo vertebral dentro da normalidade em 97,8%.

Para a análise FVC Maligna x FVC Benigna, o Radiologista 1 obteve o maior valor de sensibilidade de 0,886 entre os demais radiologistas e o classificador J48 obteve o maior valor de sensibilidade de 0,918. Isso significa que o classificador J48 tende a classificar corretamente uma FVC Maligna em 91,8%. Porém todos os valores de sensibilidade não apresentaram diferenças estatisticamente significantes baseado na avaliação dos intervalos de confiança. No geral, com exceção do classificador J48 e do Radiologista 1, houve uma queda nos valores de sensibilidade de FVC x Normal para FVC Maligna x FVC Benigna.

Ainda sobre a análise FVC Maligna x FVC Benigna, o Radiologista 3 obteve o maior valor de especificidade de 0,923 entre todas as classificações e *naïve Bayes* obteve o maior valor de especificidade de 0,887 entre os classificadores. Isso significa que o Radiologista 3 tende a classificar corretamente uma FVC Benigna em 92,3%. O classificador J48, que havia obtido o maior valor de sensibilidade de 0,918, obteve valor elevado de especificidade de 0,868. Apesar disso, todos os valores de especificidade não apresentaram diferenças estatisticamente significantes baseado na avaliação dos intervalos de confiança.

Sumarizando, desconsiderando os intervalos de confiança, a análise FVC x Normal obteve resultados de classificação com maiores valores estimados de sensibilidade e especificidade que na análise FVC Maligna x FVC Benigna. Para a análise FVC x Normal, as classificações que obtiveram ambos os valores estimados de sensibilidade e especificidade maiores que 0.9 foram os classificadores *naïve* Bayes, J48 e *k*-NN para $k = 7$. Nenhum dos três radiologistas obteve ambos os valores estimados de sensibilidade e especificidade maiores que 0.9. Quanto a análise FVC Maligna x FVC Benigna, as classificações que obtiveram ambos os valores estimados de sensibilidade e especificidade maiores que 0.8 foram os classificadores *naïve* Bayes, J48, *k*-NN para $k = 7$ e $k = 11$ e o Radiologista 1.

7.4. Comparação dos resultados com a literatura

Ribeiro et al. (2012) desenvolveram um método de detecção de FVCs da coluna lombar em radiografias laterais e obteve valores de sensibilidade de 0,78 e especificidade de 0,95. Al-Helo et al. (2013) desenvolveram um sistema CAD de detecção de fraturas vertebrais em imagens de TC e obteve valores de sensibilidade de 0,99 e especificidade de 0,875. Em nosso trabalho, o classificador *naïve* Bayes utilizando o vetor de características resultante do método automático *Wrapper* de seleção de atributos para a análise de FVC x Normal, a qual estuda a detecção de FVCs, obteve valores de sensibilidade de 0,941 com intervalo de confiança = [0,871, 0,976] e especificidade de 0,978 com intervalo de confiança = [0,914, 0,996] (Tabelas 23). Nosso método apresentou valores médios de sensibilidade 4,9 pontos percentuais abaixo que o método de Al-Helo et al. (2013), 16,1 pontos percentuais acima que o método de Ribeiro et al. (2012) e apresentou o maior valor de especificidade.

Kasai et al. (2006) com seu método de detecção de fraturas vertebrais em radiografias de tórax obteve uma acurácia de 0,766. Ghosh et al. (2011) desenvolveram um CAD completo para detecção de FVCs na coluna lombar usando imagens de TC e obteve uma acurácia de

0,973. Nosso classificador *naïve* Bayes obteve valor de acurácia de 0,958 (Tabela 20), sendo 1,53 pontos percentuais abaixo do método de Ghosh et al. (2011) e 19,2 pontos percentuais acima do método de Kasai et al. (2006).

7.5. Limitações do trabalho

Uma das limitações de nosso estudo é depender da segmentação manual, o que do ponto de vista de um CAD não é interessante. Benjelloun e Mahmoudi (2008) propuseram uma ferramenta para segmentação a qual era capaz de detectar contornos de vertebrae cervicais em imagens de raio-X utilizando um modelo de contorno dinâmico discreto e um processo de ajuste de *template* associado com um sistema de assinatura polar. Wang et al. (2015) propuseram um método de segmentação automático para vértebras e disco intervertebrais utilizando uma abordagem de segmentação por regressão nos planos axial e sagital de TC e RM. Duarte Barbieri et al. (2015) propuseram um método para segmentação automática de corpos vertebrais em IRM utilizando uma técnica de *superpixels*. Tang e Pauli (2011) desenvolveram um método totalmente automático de extração da curvatura da coluna vertebral em IRM primeiramente extraíndo as posições dos discos intervertebrais e depois utilizam essas posições para segmentar a coluna vertebral utilizando um método para o registro das vértebras. Eles observaram que uma limitação de seu método é a preocupação de que a intensidade de sinal dos discos intervertebrais deve ser diferente da intensidade de sinal dos corpos vertebrais. Com a inclusão de métodos automáticos ou semiautomáticos de segmentação, nosso método proposto para análise e classificação de corpos vertebrais poderia levar à realização de um sistema CAD clinicamente útil.

A análise de ROIs contendo apenas os corpos vertebrais pode fazer com que informações importantes fora de uma ROI, como a presença de massa epidural, intensidade de sinal do pedículo em FVCs malignas (JUNG et al., 2003) e a presença de potenciais nódulos malignos em outros corpos vertebrais no mesmo exame, não sejam observadas. É provável

que o desempenho das classificações dos Radiologistas 1, 2 e 3 fosse melhor caso eles analisassem o exame de RM e não apenas as ROIs individualmente.

O uso apenas do plano sagital mediano é outra limitação de nosso estudo, pois regiões mais laterais dos corpos vertebrais não serão analisadas e, portanto informações importantes podem ser perdidas. A extensão da segmentação e de métodos de extração de características para três dimensões seria importante para o desenvolvimento de um sistema CAD robusto para FVCs.

A análise de IRM com ponderações T2 e T1 com contraste de gadolínio poderia ser importante na detecção de possíveis sinais que poderia auxiliar na classificação de FVCs, como um corpo vertebral com sinal isointenso em imagens de ponderação T2 e T1 com contraste de gadolínio para FVCs benignas e, sinais heterogêneos ou de alta intensidade em ponderação T2 e em T1 pós-contraste de gadolínio em FVCs malignas (LINK *et al.*, 2005).

8. CONCLUSÃO

Em geral, os valores de A_z dos atributos de textura e níveis de cinza com p-valor $< 0,001$ são maiores que os valores de A_z dos atributos de forma para a análise de classificação FVC Maligna x FVC Benigna, e o oposto ocorre para a análise de FVC x Normal. Porém atributos de forma como FDF e atributos de níveis de cinza como cv não seguem essa tendência.

O método automático *Wrapper* de seleção de atributos obteve os melhores resultados para as classificações mostrando que atributos de textura e análise de níveis de cinza somados a atributos de forma melhoraram os resultados para a análise FVC x Normal, e que atributos de forma somados a atributos de textura e análise de níveis de cinza melhoraram os resultados para a análise de classificação FVC Maligna x FVC Benigna. Além disso, o método *Wrapper* diminuiu a quantidade de casos de FVCs benignas classificadas incorretamente como FVCs malignas para a análise FVC Maligna x FVC Benigna x Normal.

Nosso método obteve ótimos resultados de classificação e detecção de FVCs, e são estatisticamente comparáveis às classificações realizadas por médicos residentes em radiologista e a classificação realizada por um radiologista generalista para as mesmas ROIs. O método proposto é promissor para o desenvolvimento de um sistema CAD para a detecção de FVCs e classificação dessas FVCs em malignas ou benignas.

9. REFERÊNCIAS

AHA, D. W.; KIBLER, D.; ALBERT, M. K. Instance-based learning algorithms. **Machine Learning**, v. 6, n. 1, p. 37-67, 1991. ISSN 0885-6125.

AL-HELO, S. et al. Compression fracture diagnosis in lumbar: a clinical CAD system. **International Journal of Computer Assisted Radiology and Surgery**, v. 8, n. 3, p. 461-470, 2013. ISSN 1861-6410.

AREVALO-PEREZ, J. et al. Differentiating benign from malignant vertebral fractures using T1-weighted dynamic contrast-enhanced MRI. **Journal of Magnetic Resonance Imaging**, v. 42, n. 4, 2015. p. 1039-1048. ISSN 1053-1807.

AZEVEDO-MARQUES, P. M. et al. Classification of vertebral compression fractures in magnetic resonance images using spectral and fractal analysis. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milano, Italy, 2015. p.1-4.

BENJELLOUN, M.; MAHMOUDI, S. X-ray image segmentation for vertebral mobility analysis. **International Journal of Computer Assisted Radiology and Surgery**, v. 2, n. 6, p. 371-383, 2008. ISSN 1861-6410. Disponível em: <<http://dx.doi.org/10.1007/s11548-008-0149-1>>.

BURGES, C. A tutorial on Support Vector Machines for pattern recognition. **Data Mining and Knowledge Discovery**, v. 2, n. 2, p. 121-167, 1998. ISSN 1384-5810.

CABRAL, T. M.; RANGAYYAN, R. M. **Fractal Analysis of Breast Masses in Mammograms**. Morgan & Claypool Publishers, 2012. ISBN 1930-0328.

CUENOD, C. A. et al. Acute vertebral collapse due to osteoporosis or malignancy: appearance on unenhanced and gadolinium-enhanced MR images. **Radiology**, v. 199, n. 2, p. 541-550, 1996. ISSN 0033-8419.

DESERNO, T. M. **Biomedical Image Processing**. Aachen, Germany: Springer Science & Business Media, 2011.

DUARTE BARBIERI, P. et al. Vertebral body segmentation of spine MR images using superpixels. IEEE 28th International Symposium on Computer-Based Medical Systems (CBMS), São Carlos and Ribeirão Preto, Brazil, 2015. p.1-6.

FRIGHETTO-PEREIRA, L. et al. Semiautomatic classification of benign versus malignant vertebral compression fractures using texture and gray-level features in magnetic resonance images. *IEEE 28th International Symposium on Computer-Based Medical Systems (CBMS)*, June 2015, São Carlos and Ribeirão Preto, SP, Brazil. IEEE, 2015. p.1-5.

FRIGHETTO-PEREIRA, L. et al. Classification of vertebral compression fractures in magnetic resonance images using shape analysis. *Proceedings of IEEE E-Health and Bioengineering (EHB)*, Iasi, Romania, 2015. p.1-4.

GENANT, H. K. et al. Vertebral fracture assessment using a semiquantitative technique. **Journal of Bone and Mineral Research**, v. 8, n. 9, p. 1137-1149, Sep 1993. ISSN 0884-0431.

GHOSH, S. et al. Automatic lumbar vertebra segmentation from clinical CT for wedge compression fracture diagnosis. *SPIE Medical Imaging Proceedings International Society for Optics and Photonics*, 2011. p.1-9.

GRADOS, F. et al. Radiographic methods for evaluating osteoporotic vertebral fractures. **Joint Bone Spine**, v. 76, n. 3, p. 241-247, 2009. ISSN 1297-319x.

GRANLUND, G. H. Fourier preprocessing for hand print character recognition. **IEEE Transactions on Computers**, v. C-21, n. 2, p. 195-202, 1972. ISSN 0018-9340.

GUGLIELMI, G. et al. Assessment of osteoporotic vertebral fractures using specialized workflow software for 6-point morphometry. **European Journal of Radiology**, v. 70, n. 1, p. 142-149, 2009. ISSN 0720-048x.

HALL, M. et al. The WEKA data mining software: an update. **ACM SIGKDD explorations newsletter**, v. 11, n. 1, p. 10-19, 2009. ISSN 1931-0145.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. H. Textural features for image classification. **IEEE Transactions on Systems, Man, and Cybernetics**, n. 6, p. 610-632, 1973. ISSN 0018-9472.

HU, M. K. Visual pattern recognition by moment invariants. **IRE Transactions on Information Theory**, v. 8, n. 2, p. 179-188, 1962. ISSN 0096-1000.

JOANES, D. N.; GILL, C. A. Comparing measures of sample skewness and kurtosis. **Journal of the Royal Statistical Society**, v. 47, n. 1, p. 183-190, 1998.

JOHN, G. H.; LANGLEY, P. Estimating continuous distributions in Bayesian classifiers. 11th Conference on Uncertainty in Artificial Intelligence, Montreal, Canada. Morgan Kaufmann Publishers Inc., 1995. p.1-8.

JUNG, H. S. et al. Discrimination of metastatic from acute osteoporotic compression spinal fractures with MR imaging. **Radiographics**, v. 23, n. 1, p. 179-188, 2003. ISSN 0271-5333.

KAPLAN, P. A. et al. **Musculoskeletal MRI**. Philadelphia: Saunders Elsevier, 2001.

KASAI, S. et al. Computerized detection of vertebral compression fractures on lateral chest radiographs: preliminary results with a tool for early detection of osteoporosis. **Medical Physics**, v. 33, n. 12, p. 4664-4675, 2006. ISSN 0094-2405.

KHAN, A. A. et al. Principal component and factor analysis to study variations in the aging lumbar spine. **IEEE Journal of Biomedical and Health Informatics**, v. 19, n. 2, p. 745-751, 2015. ISSN 2168-2194.

KOHAVI, R.; JOHN, G. H. Wrappers for feature subset selection. **Artificial Intelligence**, v. 97, n. 1, p. 273-325, 1997. ISSN 0004-3702.

KONDO, K. L. Osteoporotic vertebral compression fractures and vertebral augmentation. **Seminars in Interventional Radiology**, v. 25, n. 4, p. 413-425, 2008. ISSN 0739-9529. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/21326583>>.

LEE, C.-C. et al. Robust radial basis function neural networks. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 29, n. 6, p. 674-686, 1999. ISSN 1083-4419.

LINK, T. M. et al. Radiologic assessment of osteoporotic vertebral fractures: diagnostic and prognostic implications. **European Radiology**, v. 15, n. 8, p. 1521-1532, 2005. ISSN 0938-7994.

MASSEY JR, F. J. The Kolmogorov-Smirnov test for goodness of fit. **Journal of the American Statistical Association**, v. 46, n. 253, p. 68-79, 1951. ISSN 0162-1459.

MATLAB. Mathworks Incorporated 2014.

METZ, C. E. **Rockit**: University of Chicago 1998.

OEI, L. et al. Review of radiological scoring methods of osteoporotic vertebral fractures for clinical and research settings. **European Radiology**, v. 23, n. 2, p. 476-487, 2013. ISSN 1432-1084. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/22892721>>.

PERSOON, E.; FU, K.-S. Shape discrimination using Fourier descriptors. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 7, n. 3, p. 170-180, 1977. ISSN 0018-9472.

Photoshop. Adobe Systems Incorporated 2010.

PRASAD, D.; SCHIFF, D. Malignant spinal-cord compression. **The Lancet Oncology**, v. 6, n. 1, p. 15-25, Jan 2005. ISSN 1470-2045.

QUILAN, J. R. **C4.5: Programs for Machine Learning**. San Mateo: Morgan Kaufmann, 1993. ISBN 1-55860-238-0.

RAGUSO, G. et al. Application of fractal analysis to mammography. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Buenos Aires, Argentina, 2010. p.1-4.

RANGAYYAN, R. M. **Biomedical Image Analysis**. CRC Press, 2005. Disponível em: <http://books.google.com.br/books?id=o_8jEGNpEuwC>.

RIBEIRO, E. A. et al. Detection of vertebral compression fractures in lateral lumbar X-ray images. XXIII Congresso Brasileiro em Engenharia Biomédica (CBEB), Porto de Galinhas, Brazil, 2012. p.1-4.

RISH, I. An empirical study of the naive Bayes classifier. **IJCAI workshop on empirical methods in artificial intelligence**, v. 3, n. 22, p. 41-46, 2001.

ROSENFELD, A.; KAK, A. C. **Digital picture processing**: Elsevier, 2014.

SEMELKA, R. C. et al. Imaging strategies to reduce the risk of radiation in CT studies, including selective substitution with MRI. **Journal of Magnetic Resonance Imaging**, v. 25, n. 5, p. 900-910, 2007. ISSN 1053-1807.

SHEN, L.; RANGAYYAN, R. M.; DESAUTELS, J. L. Application of shape analysis to mammographic calcifications. **IEEE Transactions on Medical Imaging**, v. 13, n. 2, p. 263-274, 1994.

SOILLE, P. **Morphological Image Analysis: Principles and Applications**. 2. Berlin Heidelberg: Springer Science & Business Media, 2013.

TANG, Z.; PAULI, J. Fully automatic extraction of human spine curve from MR images using methods of efficient intervertebral disk extraction and vertebra registration. **International Journal of Computer Assisted Radiology and Surgery**, v. 6, n. 1, p. 21-33, 2011. ISSN 1861-6410.

TEHRANZADEH, J.; TAO, C. Advances in MR imaging of vertebral collapse. **Seminars in Ultrasound, CT, and MR**, v. 25, n. 6, p. 440-461, 2004. ISSN 0887-2171. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/15663316>>.

UETANI, M.; HASHMI, R.; HAYASHI, K. Malignant and benign compression fractures: differentiation and diagnostic pitfalls on MRI. **Clinical Radiology**, v. 59, n. 2, p. 124-132, 2004. ISSN 0009-9260.

WANG, Z. et al. Regression segmentation for M³ spinal images. **IEEE Transactions on Medical Imaging**, v. 34, n. 8, p. 1640-1649, 2015. ISSN 0278-0062.

WILCOXON, F.; KATTI, S. K.; WILCOX, R. A. Critical values and probability levels for the Wilcoxon rank sum test and the Wilcoxon signed rank test. **Selected Tables in Mathematical Statistics**, v. 1, p. 171-260, 1970.

ZAHN, C. T.; ROSKIES, R. Z. Fourier descriptors for plane closed curves. **IEEE Transactions on Computers**, v. 10, n. 3, p. 269-282, 1972. ISSN 0018-9340.

Anexo A

HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA
DE RIBEIRÃO PRETO DA UNIVERSIDADE DE SÃO PAULO

Ribeirão Preto, 25 de fevereiro de 2015

Ofício nº 737/2015
CEP/MGV

Prezados Senhores,

O trabalho intitulado “**CLASSIFICAÇÃO SEMIAUTOMÁTICA DE FRATURAS VERTEBRAIS BENIGNAS E MALIGNAS EM IMAGENS DE RESSONÂNCIA MAGNÉTICA**”, foi analisado pelo Comitê de Ética em Pesquisa, em sua 402ª Reunião Ordinária realizada em 23/02/2015 e enquadrado na categoria: APROVADO, bem como a solicitação de dispensa do Termo de Consentimento Livre e Esclarecido, de acordo com o Processo HCRP nº 13414/2014.

Este Comitê segue integralmente a Conferência Internacional de Harmonização de Boas Práticas Clínicas (IGH-GCP), bem como a Resolução nº 466/12 CNS/MS.

Lembramos que devem ser apresentados a este CEP, o Relatório Parcial e o Relatório Final da pesquisa.

Atenciosamente.

DRª. MARCIA GUIMARÃES VILLANOVA
Coordenadora do Comitê de Ética em
Pesquisa do HCRP e da FMRP-USP

Ilustríssimos Senhores

LUCAS FRIGHETTO PEREIRA
PROF.DR. MARCELLO HENRIQUE NOGUEIRA BARBOSA(Orientador)
Depto. de Clínica Médica

Anexo B

Computers in Biology and Medicine 73 (2016) 147–156

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Shape, texture and statistical features for classification of benign and malignant vertebral compression fractures in magnetic resonance images

Lucas Frighetto-Pereira^a, Rangaraj Mandayam Rangayyan^b, Guilherme Augusto Metzner^a, Paulo Mazzonecini de Azevedo-Marques^a, Marcello Henrique Nogueira-Barbosa^{a,*}

^a Image Science and Medical Physics Center, Internal Medicine Department, Ribeirão Preto Medical School, University of São Paulo, 3900 Bandeirantes Avenue, Ribeirão Preto, SP 14048-900, Brazil

^b Department of Electrical and Computer Engineering, Schulich School of Engineering, University of Calgary, 2500 University Dr. NW, Calgary, AB, Canada T2N 1N4

ARTICLE INFO

Article history:

Received 11 January 2016

Received in revised form

8 April 2016

Accepted 9 April 2016

Keywords:

Computer-aided diagnosis

Image processing

Magnetic resonance images

Shape analysis

Statistical analysis of gray levels

Texture analysis

Vertebral compression fractures

ABSTRACT

Purpose: Vertebral compression fractures (VCFs) result in partial collapse of vertebral bodies. They usually are nontraumatic or occur with low-energy trauma in the elderly secondary to different etiologies, such as insufficiency fractures of bone fragility in osteoporosis (benign fractures) or vertebral metastasis (malignant fractures). Our study aims to classify VCFs in T1-weighted magnetic resonance images (MRI).

Methods: We used the median sagittal planes of lumbar spine MRIs from 63 patients (38 women and 25 men) previously diagnosed with VCFs. The lumbar vertebral bodies were manually segmented and statistical features of gray levels were computed from the histogram. We also extracted texture and shape features to analyze the contours of the vertebral bodies. In total, 102 lumbar VCFs (53 benign and 49 malignant) and 89 normal lumbar vertebral bodies were analyzed. The k-nearest-neighbor method, a neural network with radial basis functions, and a naïve Bayes classifier were used with feature selection. We compared the classification obtained by these classifiers with the final diagnosis of each case, including biopsy for the malignant fractures and clinical and laboratory follow up for the benign fractures.

Results: The results obtained show an area under the receiver operating characteristic curve of 0.97 in distinguishing between normal and fractured vertebral bodies, and 0.92 in discriminating between benign and malignant fractures.

Conclusions: The proposed classification methods based on shape, texture, and statistical features have provided high accuracy and may assist in the diagnosis of VCFs.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Vertebral compression fracture (VCF) is a fracture with partial collapse of the affected vertebral body. Clinically, a radiologist will have no doubt about the etiology of a traumatic VCF as a patient typically presents with a recent history of trauma, such as a fall or an accident. However, when the patient develops a recent vertebral collapse without trauma, or when there was a low-energy traumatic incident, the etiology needs to be investigated, and there can be difficulty in arriving at a diagnosis. This problem is common especially in the elderly population due to the higher incidence of VCFs secondary to bone failure. Vertebral fragility

fractures secondary to osteoporosis are an increasingly important health issue [1,2]. The elderly population also has a high incidence of VCFs related to metastatic cancer affecting bone [3]. Clinically, a VCF secondary to osteoporosis is labeled as a benign VCF, and a VCF caused by bone metastasis is defined as a malignant VCF [3]. Complaints from patients related to both types of VCFs, benign and malignant, may be similar; however, the differentiation between the two types of VCFs is fundamental for correct diagnosis and appropriate treatment.

Magnetic resonance imaging (MRI) is effective in early detection of vertebral fractures [4,5]. The distribution of the signal intensity abnormality throughout the vertebral body is an important criterion to discriminate between benign and malignant VCFs in MRI, but the analysis of vertebral body shape can also help in this process [6–9]. A malignant vertebral collapse typically exhibits diffuse low signal intensity throughout the vertebral body in T1-weighted MRI.

* Corresponding author.

E-mail address: marcello@fmrp.usp.br (M.H. Nogueira-Barbosa).

<http://dx.doi.org/10.1016/j.combiomed.2016.04.006>

0010-4825/© 2016 Elsevier Ltd. All rights reserved.

Anexo C

Int J CARS (2016) 11 (Suppl 1):S1–S286

S123

Conclusion

This is the first trial and proof-of concept study to demonstrate equivalent specificity and non-inferior or even superior sensitivity of the e-ASPECTS software compared to ASPECTS scoring by board-certified neuroradiologists (with specific expertise in acute stroke imaging) of baseline NCCTs from acute ischemic stroke patients. A previous monocentric trial already demonstrated superiority of e-ASPECTS to stroke trainees regarding sensitivity of a region based analysis approach [3].

The strengths of this study are the multicentric design, which meant that NCCTs were acquired with various CT scanners indicating generalisation of the results. A limitation is surely that the ground truth was based on the follow up CT scan and not on expeditious MRI; however that should not hamper the overall comparison between humans and e-ASPECTS. Although e-ASPECTS is now CE-marked, careful assessment of each CT scan by the physician to rule out other pathologies and haemorrhage is still mandatory. It is not intended to replace the physician's assessment of the scan but it can be a valuable second opinion and confirmation of the own interpretation on expert level.

e-ASPECTS might be a valuable tool especially for hospitals, where expert assessment of NCCT imaging is not available 24/7.

References

- [1] Barber PA, Demchuk AM, Zhang J, Buchan AM (2000) Validity and reliability of a quantitative computed tomography score in predicting outcome of hyperacute stroke before thrombolytic therapy. Aspects study group. *Alberta stroke programme early ct score*. *Lancet*;355:1670–1674
- [2] Hampton-Till J, Harrison M, Kühn AL, Anderson O, Sinha D, Tysoe S, et al. (2015) Automated quantification of stroke damage on brain computed tomography scans: E-aspects. *European Medical Journal* 3[1]:69–74
- [3] Herweh C, Ringleb PA, Rauch G, Geny S, Behrens L, Möhlenbruch M, et al. (2015) Performance of e-ASPECTS software in comparison to that of stroke physicians on assessing CT scans of acute ischemic stroke patients. *International Journal of Stroke* (in press, accepted 12.Nov. 2015)

Recognition of vertebral compression fractures in magnetic resonance images using measures of distances to the centroid

L. Frigetto-Pereira¹, R. M. Rangayyan², G. A. Metzner¹,

P. M. Azevedo-Marques¹, M. H. Nogueira-Barbosa¹

¹University of São Paulo, Ribeirão Preto Medical School, Ribeirão Preto, Brazil

²University of Calgary, Department of Electrical and Computer Engineering, Schulich School of Engineering, Calgary, Canada

Keywords Magnetic resonance imaging · Distances to the centroid · Shape analysis · Vertebral compression fracture

Purpose

A common clinical problem in the lumbar spine is vertebral compression fracture (VCF), particularly in elderly patients. A benign VCF is a consequence of osteoporotic fragility and is the most common type of non-traumatic vertebral fracture in the elderly. Bone metastatic disease may cause malignant VCF that is also common in the elderly population. The recognition of VCFs and differentiation between the two types of VCFs is fundamental for appropriate treatment [1]. Currently, magnetic resonance imaging (MRI) is the most suitable and effective imaging modality for early detection and investigation for differential diagnosis of nontraumatic vertebral fractures.

The shapes of vertebral bodies may show multiple types of changes after a compression fracture. A vertebral body without

fracture usually presents a nearly rectangular shape in the sagittal plane MRI, with possible mild concavity in the superior and inferior plateaus and in the posterior and anterior walls of the vertebral body. Benign VCFs can present accentuated concavity in the vertebral plateaus and may cause posterolateral wall fragment retropulsion with angulated or irregular contours; subchondral bone impaction can cause rough and notched contours. Malignant VCFs may also have concave deformation of the vertebral plateaus; however, their contours may be relatively rounded or smoothed due to bulging neoplastic tissue, which can also lead to convexity of the posterior vertebral body wall [1,2].

We propose image-processing methods based on measurement of the distance from a vertebral body's contour to its centroid to recognize VCFs. We applied these methods to images in the median sagittal plane of T1-weighted MRI exams.

Methods

This study has been approved by the Research Committee of Ethics of the Ribeirão Preto Medical School. We used a set of T1-weighted MRI exams of 63 patients, including 38 women and 25 men, with the mean age of 62.25 ± 14.13 years, collected from the Radiology Information System (RIS) of our university hospital. Each patient was diagnosed with at least one lumbar VCF. Manual segmentation of the lumbar vertebral bodies was performed to define the regions of interest (ROIs) for the study. The data set includes 102 lumbar VCFs (53 benign and 49 malignant) and 89 nonfractured (normal) lumbar vertebral bodies.

Euclidean distances from the centroid of a given ROI to each pixel on its boundary were computed. Figure 1 shows examples of contours of a normal vertebral body, a benign VCF, and a malignant VCF with the distances to their centroids. We computed seven features using the distances for each ROI: the mean, standard deviation, variance, coefficient of variation, skewness, kurtosis [3–5], and a combination of the fourth-order and second-order moments (m_f) as defined by Shen et al. [5].

Fig. 1 Examples of distances between a vertebral body's contour and its centroid: (a) normal vertebral body, (b) benign vertebral compression fracture (VCF), and (c) malignant VCF

We performed two types of classification: both types of VCFs together versus normal vertebral bodies and benign VCFs versus malignant VCFs using the k -nearest-neighbor (k -NN) method with $k = 5$, the naïve Bayes method, and a neural network with radial basis functions (RBF network) as classifiers. Before the classification step, feature selection was performed using the wrapper method with 10-fold cross-validation. The performance of each classification procedure was analyzed using the area under the receiver operating characteristic (ROC) curve (AUROC) and accuracy (ACC) computed by using the confusion matrix.

Results

The AUROC values for the classification of all VCFs together versus normal vertebral bodies varied from 0.93 to 0.96. The best classifier for this type of classification was the naïve Bayes, which obtained AUROC = 0.96, ACC = 89 %, sensitivity = 0.84, and specificity = 0.94. For the naïve Bayes classifier, only kurtosis was not selected in the feature selection step. A high value of specificity was also obtained in three-way classification separating malignant VCFs,

Recognition of vertebral compression fractures in magnetic resonance images using statistics of height and width

Frighetto-Pereira L, Metzner GA, Azevedo-Marques PM,
Nogueira-Barbosa MH
Image Science and Medical Physics Center,
Internal Medicine Department
Ribeirão Preto Medical School, University of São Paulo
Ribeirão Preto, São Paulo, Brazil
marcello@fmrp.usp.br

Oloumi F, Rangayyan RM
Department of Electrical and Computer Engineering
Schulich School of Engineering, University of Calgary
Calgary, Alberta, Canada

Abstract— A vertebral compression fracture (VCF) results in vertebral body partial collapse and is often non-traumatic in elderly patients occurring secondary to osteoporosis and bone fragility. Such fractures are benign and may be confounded with pathologic fractures secondary to metastatic cancer. Differentiation between the two types of VCFs is important for correct diagnosis and treatment. Our study aims to recognize and classify VCFs in T1-weighted magnetic resonance images (MRI). We used images of 102 lumbar VCFs (53 benign and 49 malignant) and 89 normal vertebral bodies manually segmented from the median sagittal planes of lumbar spine MRIs of 63 patients (38 women and 25 men). The principal axis of each vertebral body region of interest was identified using moments. Statistical features of height and width measured perpendicular and parallel to the principal axis were computed. The *k*-nearest-neighbor method, a neural network with radial basis functions, and the naïve Bayes classifier were used with feature selection for classification. Areas under the receiver operating characteristic curve of 0.96 in the recognition of VCFs as compared with normal vertebral bodies and 0.73 for the classification of benign versus malignant VCFs were obtained. The proposed methods are promising for the recognition of VCFs, but additional features are needed to improve the classification of benign versus malignant VCFs.

Keywords—magnetic resonance imaging, principal axis, shape analysis, vertebral compression fracture.

I. INTRODUCTION

A. Vertebral compression fractures and shape features

The elderly population has higher incidence and prevalence of fractures secondary to bone failure as compared with the younger population. A vertebral compression fracture (VCF) is clinically significant and evolves with partial collapse of the vertebral body. Such fractures are frequently encountered in elderly patients without history of trauma or occurring with low-energy trauma, such as a fall off the bed. The most common type of VCF is benign and occurs secondary to

osteoporosis [1-3]. The elderly population also has higher incidence and prevalence of VCFs related to metastatic cancer affecting bone, which are labeled as malignant VCFs [3]. Benign and malignant VCFs can generate similar complaints from patients; however, differentiation between benign and malignant is crucial for the correct diagnosis and appropriate treatment. Magnetic resonance imaging (MRI) is the most suitable imaging modality for spinal abnormalities that is effective for detection of VCFs [4-6].

In the sagittal plane, normal vertebral bodies usually present nearly rectangular shapes. However, they could include marginal bone proliferation and small bone outgrowths known as osteophytes, which may cause shape abnormalities. VCFs usually cause several vertebral bodies shape abnormalities. A malignant VCF could lead to convexity of the posterior vertebral body wall due to bulging of neoplastic tissue and may cause the contour of the vertebra to be relatively rounded or smoothed. On the other hand, a benign VCF may cause posteriorsuperior wall fragment retropulsion and an irregular or angulated contour. The vertebral plateaus may have a more accentuated concave shape and subchondral bone impaction may lead to rough contours with indentations [7, 8].

B. Review of studies on VCFs

Genant et al. [9] proposed a grading method for semiquantitative evaluation of osteoporotic vertebral fractures. This method compares the heights of the anterior, central, and posterior portions of vertebral bodies. Normal vertebral bodies are given grade zero. Mild deformity, with 20–25% of reduction in vertebral height in any part, is defined as grade one. Grade two is given for moderate deformity with a reduction of vertebral height in any part in the range 25–40%. A reduction of more than 40% of vertebral height in any part is considered severe and grade three.

Guglielmi et al. [10] evaluated a software tool for model-based image analysis using a 6-point morphometry protocol,

Classification of Vertebral Compression Fractures in Magnetic Resonance Images using Shape Analysis

L. Frighetto-Pereira¹, R. Menezes-Reis¹, G. A. Metzner¹, R. M. Rangayyan², Fellow, IEEE,

M. H. Nogueira-Barbosa¹, P. M. Azevedo-Marques¹, Member, IEEE.

¹Ribeirão Preto Medical School, University of São Paulo, USP, Ribeirão Preto, Brazil.

²Schulich School of Engineering University of Calgary, Calgary-AB, Canada.

Abstract- Fractures with partial collapse of vertebral bodies are referred to as "vertebral compression fractures" or VCFs. VCFs can have different etiologies comprising trauma, bone failure related to osteoporosis (benign), or metastatic cancer (malignant) affecting bone. This paper presents a study aimed to help in the classification of VCFs in magnetic resonance imaging (MRI). We used T1-weighted MRI of the lumbar spine in the sagittal plane. Images from 63 patients (38 women, 25 men) diagnosed with at least one VCF were studied. Shape features as normalized compactness, convex deficiency, Hu's moments and Fourier-descriptor-based feature were extracted from manually segmented images of 102 lumbar vertebral bodies with VCFs (53 benign and 49 malignant), and 106 normal lumbar vertebral bodies. The classification was performed using different classifiers, as k-nearest-neighbor, naïve Bayesian classifier and neural network with radial basis functions with different feature combinations. Results obtained show that shape analysis was able to correctly distinguish between normal and fractured vertebral bodies with a correct classification rate up to 90.58% and between benign and malignant with a correct classification rate up to 80.39%.

Keywords: Vertebral Compression Fractures, Magnetic Resonance Images, Shape Analysis.

I. INTRODUCTION

Osteoporosis is the most common metabolic bone disease, and vertebral compression fractures (VCFs) are the most common type of osteoporotic fractures. Such fractures are associated with significant morbidity, mortality, and high health-care costs. Given the aging of populations, osteoporotic vertebral fractures are likely to become an increasingly important health issue [1], [2]. On the other hand, the elderly population has also a high incidence and prevalence of VCFs related to metastatic cancer affecting bone [3]. Clinically, a VCF secondary to osteoporosis may be labeled as a benign fracture whereas a VCF caused by bone metastasis is usually defined as a malignant fracture [3], [4]. Although complaints related to both types of VCFs, benign and malignant, may be similar, the differentiation between these two types of vertebral collapse is fundamental for the diagnosis and treatment of the patient [4].

At present, magnetic resonance imaging (MRI) is considered to be the most reliable and most widely used imaging method for spine diseases, and it is effective in early detection of fractures [5], [6]. The distribution of the signal intensity abnormality throughout the vertebral body is an important criterion to help in the differentiation between a benign and a malignant VCF in MRI, but vertebral body shape may also help in this differentiation [7]. VCFs cause multiple types of changes in the shapes of vertebrae.

Fig. 1. Examples of spine T1-weighted MRI used in the study: (A) Normal exam. (B) Benign VCFs in L1 and L4. (C) Malignant VCFs in L1 and L5.

Fig. 2. Three examples of each kind of vertebral body ROIs

Malignant VCFs are known to cause posterior vertebral body cortex convex bulging more frequently while benign VCF may cause a posterior wall fragment retropulsion with angulated or irregular contours [7].

In the case of benign VCFs, the vertebral plateaus may acquire a more accentuated concave shape, and subchondral bone impaction may result in rough contours with indentations. Malignant VCFs could result in a posterior bulge or convexity of the posterior vertebral body wall though it may occur together with a concave deformation in the region of the vertebral plateaus. Malignant processes may cause the contours of vertebrae to be relatively rounded or smoothed due to bulging neoplastic tissue. Figure 1 shows parts of sagittal MRI sections of three subjects with normal vertebrae, benign VCFs, and malignant VCFs, where the

Anexo F

Semiautomatic classification of benign versus malignant vertebral compression fractures using texture and gray-level features in magnetic resonance images

L. Frighetto-Pereira¹, R. Menezes-Reis¹, G. A. Metzner¹, R. M. Rangayyan², *Fellow, IEEE*,
 P. M. Azevedo-Marques¹, *Member, IEEE*, M. H. Nogueira-Barbosa¹.
 1- Ribeirão Preto Medical School, University of São Paulo, USP, Ribeirão Preto, Brazil.
 2- Schulich School of Engineering University of Calgary, Calgary-AB, Canada.
 e-mail: lfplucas@usp.br

Abstract—Our study aimed to develop a system for computer-aided diagnosis of vertebral compression fractures (VCFs) using magnetic resonance imaging (MRI), to help in the differentiation between malignant and benign VCFs. Lumbar spine MRI was used to acquire T1-weighted images in the sagittal plane. Images from 63 consecutive patients (38 women, 25 men, mean age 62.25 ± 14.13 years) with at least one VCF diagnosis were studied. Contrast and texture features were extracted from manually segmented images of 103 vertebral bodies with VCFs. The classification of malignant vs. benign VCFs was performed using the k-nearest neighbor (KNN) classifier with the Euclidean distance. Using a KNN classifier with $k=3$, feature selection, and 10-fold cross-validation, we obtained a value of the area under the receiver operating characteristic curve of 0.913.

Keywords: Compression Fracture; Lumbar Vertebra; Osteoporosis; Metastasis; Magnetic Resonance Imaging; Image Processing.

I. INTRODUCTION

Fractures with partial collapse of vertebral bodies are generically named "vertebral compression fractures" (VCFs). In the clinical setting, traumatic VCFs raise no doubt about their etiology because, typically, the radiologist will find an acute vertebral collapse and there will be a recent history of trauma, such as a fall over stairs or a car accident. The difficulty in defining the cause of a VCF arises when the patient has developed a recent vertebral collapse and there is no history of trauma, or even when there was a low-energy trauma, such as a fall from the bed. This diagnostic dilemma is common especially in the elderly. The elderly population has higher incidence and prevalence of VCFs secondary to bone failure than the younger population. Such bone failure may be related to osteoporosis or bone metastasis [1]. Osteoporotic VCFs are an important risk factor for additional future osteoporotic fractures and are also linked to increased mortality [2]. VCFs secondary to bone metastases may represent the onset of symptoms of a patient with neoplasia or may represent a complication in patients already diagnosed with this systemic disease. Clinically, a VCF secondary to osteoporosis may be labeled as a benign fracture whereas a VCF caused by bone

metastasis is usually defined as a malignant fracture [3] [4].

Complaints related to both types of VCFs, benign and malignant, may be similar, but the differentiation between these two types of vertebral collapse is fundamental for the diagnosis and treatment of the patient [3].

Images of malignant vertebral collapse typically exhibit low signal intensity throughout the whole vertebral body involved in T1-weighted magnetic resonance imaging (MRI). Osteoporotic VCFs characteristically demonstrate partial preservation of the normal fatty bone marrow signal in the vertebral body on T1-weighted MRI [5]. Figure 1 shows examples of sagittal T1-weighted spine MRI acquired from different patients to illustrate some of the VCF patterns identified in the MRI dataset used in the present study.

Since not all cases are typically benign or malignant VCFs in clinical practice, it is easy to recognize the potential positive impact of the development of accurate and reproducible diagnostic methods for VCFs, with the aid of computational techniques.

Kasai et al. [6] described an automatic computer-aided diagnosis (CAD) system for vertebral fractures based on X-ray images of the chest. They obtained accuracy ranging between 70.9% and 76.6% in the detection of vertebral fracture compared with the identification of vertebral fractures based on manual measurements performed by radiologists.

Ribeiro et al. [7] proposed a CAD method to facilitate the detection of VCFs in lateral radiographs of the lumbar spine with a sensitivity of up to 78% and specificity of 95%.

Al-Helo et al. [8] developed a CAD system to detect vertebral body fractures in computed tomographic (CT) images and obtained a specificity of 87.5% and sensitivity over 99% using K-Means. They performed vertebrae localization and segmentation, and then classified each vertebra as fractured or non-fractured, based on an Active Shape Model and a Gradient Vector Flow Active Contours.

The studies reviewed above have in common the development of CAD systems to detect vertebral fractures using imaging methods based on ionizing radiation, using planar radiographs or CT scans. Currently, there is a recognized need to reduce the dose of ionizing radiation resulting to patients from medical imaging procedures [9].

Anexo G

XXIV Congresso Brasileiro de Engenharia Biomédica – CBEB 2014

CLASSIFICAÇÃO SEMIAUTOMÁTICA DE FRATURAS VERTEBRAIS POR COMPRESSÃO EM IMAGENS DE RESSONÂNCIA MAGNÉTICA

L. Frighetto-Pereira*, R. M. Reis*, L. C. Pereyra*, P. M. Azevedo-Marques***, R. M. Rangayyan**, M. H. Nogueira-Barbosa*.

* Faculdade de Medicina de Ribeirão Preto - Universidade de São Paulo, Ribeirão Preto - SP, Brasil.

** Schulich School of Engineering, University of Calgary, Calgary-AB, Canada.

e-mail: lfplucas@usp.br

Resumo: Esse artigo apresenta o uso de técnicas de processamento de imagens para Diagnóstico Auxiliado por Computador (CAD) de fraturas vertebrais lombares por compressão (VCFs) em imagens de ressonância magnética (RM). Foram utilizadas imagens de RM obtidas na rotina clínica, no plano sagital, ponderadas em tempo de relaxação T1. Características de forma (área do corpo vertebral), de contraste (variação dos níveis de cinza do histograma) e de textura (14 atributos de Haralick) foram extraídas de 52 corpos vertebrais segmentados manualmente. A classificação dos padrões vertebrais foi feita utilizando-se um classificador não paramétrico do tipo KNN (*k-nearest neighbors*), com distância Euclidiana. O padrão de referência utilizado para testar o desempenho da classificação semiautomática foi uma classificação da presença ou ausência de fratura vertebral realizada por um radiologista experiente. Considerando um esquema de treinamento e testes com validação cruzada e dez repetições, obteve-se uma classificação correta de fratura vertebral de 82,69%.

Palavras-chave: corpos vertebrais, fraturas, ressonância magnética, reconhecimento de padrões.

Abstract: This paper presents the use of image processing techniques for Computer Aided Diagnosis (CAD) of vertebral compression fractures (VCFs) in magnetic resonance imaging (MRI). MR images obtained in the clinical routine in the sagittal plane, weighted relaxation time T1 have been used. Shape characteristics (area of the vertebral body), contrast (varying levels of gray histogram) and texture (Haralick 14 attributes) have been extracted from 52 manually segmented vertebral bodies. Classification of vertebral patterns was performed using a non-parametric classifier type KNN (*k-nearest neighbors*), with Euclidean Distance. The benchmark used to test the performance of semi-automatic classification was a classification of the presence or absence of vertebral fracture performed by an experienced radiologist. Considering a system of training and testing with cross-validation and ten repetitions, we obtained a correct classification of vertebral fracture of 82.69%.

Keywords: vertebral bodies, fractures, magnetic resonance, pattern recognition.

Introdução

Fraturas compressivas dos corpos vertebrais não traumáticas em geral são secundárias às doenças que causam fragilidade óssea como a osteoporose e as metástases ósseas. Tanto as fraturas secundárias à osteoporose quanto as fraturas patológicas decorrentes de neoplasias malignas apresentam significativo impacto na qualidade de vida, no tratamento dos pacientes, e nos custos para o sistema de saúde, sendo particularmente comuns em idades mais avançadas.

Fraturas vertebrais compressivas secundárias às insuficiências ósseas da osteoporose frequentemente ocorrem de forma subclínica. De qualquer maneira as fraturas vertebrais secundárias a osteoporose constituem importante fator de predição de futuras fraturas da osteoporose e também estão relacionadas a aumento da mortalidade. As fraturas secundárias a metástases por sua vez podem representar o início dos sintomas de um paciente com neoplasia ou podem representar uma complicação em pacientes já diagnosticados da doença sistêmica. Qualquer que seja a etiologia das fraturas vertebrais compressivas a identificação precisa destas alterações é fundamental para o diagnóstico e para o tratamento dos pacientes. A Figura 1 ilustra uma imagem de ressonância magnética com ponderação em T1 da coluna vertebral contendo uma fratura vertebral em L1.

Desta forma é fácil entender o potencial impacto positivo do desenvolvimento de métodos diagnósticos precisos e reprodutíveis para as fraturas compressivas vertebrais, com o auxílio de técnicas computacionais.

KASAI *et al*, 2006 [1] descreveram um sistema automático de diagnóstico auxiliado por computador (CAD) para fraturas vertebrais a partir de imagens de radiografias em perfil do tórax. Aqueles autores obtiveram acurácia variando entre 70,9% e 76,6% quando comparado com a identificação de fraturas vertebrais baseada nas marcações manuais realizadas por radiologistas. RIBEIRO *et al*, 2012 [2] propuseram um CAD para auxílio na detecção de fraturas por compressão dos corpos vertebrais em radiografias de perfil da coluna lombar com sensibilidade de 78% e especificidade de 95%. AL-HELO *et al*, 2013 [3] desenvolveram um CAD para a detecção de fraturas de corpos vertebrais a partir de imagens de tomografia