CLASSIFICAÇÃO SEMI-AUTOMÁTICA DA DEGENERAÇÃO DE DISCOS INTERVERTEBRAIS EM IMAGENS DE RESSONÂNCIA MAGNÉTICA

M. S. Barreiro* **, M. H. Nogueira-Barbosa**, L. G. Savarese**,
R. M. Reis** e P. M. Azevedo-Marques**

*Departamento de Engenharia de Computação, IFTM, Uberaba/MG, Brasil
**Departamento de Clínica Médica, USP, Ribeirão Preto/SP, Brasil

Resumo: Este trabalho descreve o desenvolvimento de um método que classifica de forma semi-automática a degeneração de discos intervertebrais lombares em imagens de ressonância magnética ponderadas em T2. O conjunto de imagens consiste de 210 discos extraídos de exames de 94 indivíduos (20 a 80 anos). A classificação é feita por uma rede neural do tipo perceptron multicamada com 6 entradas, 15 neurônios na camada intermediária e 1 saída. Os resultados obtidos mostraram uma taxa média de acerto de 81,42%, com erro padrão de 9,11%.

Palavras-chave: Degeneração do Disco Intervertebral, Imagens de Ressonância Magnética Ponderadas em T2, Processamento de Imagens, Análise de Textura, Classificação de Pfirrmann.

Abstract: *This article describes the development of a method that classifies semi-automatic degeneration of lumbar intervertebral discs in magnetic resonance T2-weighted images. The dataset consists of images of 210 discs obtained from 94 individuals (20 to 80 year old). An artificial neural network of the multilayer perceptron with 6 inputs, 15 neurons in the hidden layer and 1 output, was used to check the efficiency of this study. Obtained an average rate of sucess of 81.42%, with a standard error of 9.11%.*

Keywords: *Intervertebral Disc Degeneration, Weighted Magnetic Resonance Images T2, Image Processing, Texture Analysis, Pfirrmann's Scale.*

Introdução

A degeneração do disco intervertebral afeta milhões de pessoas em todo o mundo, e isto pode ocorrer por diversos fatores, como tabagismo, obesidade, trauma, excesso de atividade física ou sedentarismo (1).

Atualmente, a Ressonância Magnética (RM) é considerada o método não-invasivo mais sensível na avaliação por imagem do disco intervertebral e das estruturas dos sistemas neuromusculoesqueléticos adjacentes, quando da avaliação da existência de degeneração discal (2).

Existem técnicas quantitativas específicas em RM para avaliação do disco intervertebral, sendo que a maioria dos sistemas e escalas de classificação de sua degeneração empregados na rotina clínica se utiliza de imagens ponderadas em T2. Nas imagens ponderadas em T2, a água é caracterizada pelo sinal de alta intensidade, favorecendo a observação da perda do sinal das alterações morfológicas na degeneração do disco (3). A escala de graduação mais difundida na literatura para classificar esse nível de degeneração é a escala de Pfirrmann (4), que classifica o nível de degeneração de acordo com a intensidade do sinal do núcleo pulposo, a estrutura do disco, a altura do disco e a delimitação entre

núcleo pulposo e ânulos fibroso (5). Os níveis vão de I a V, quanto maior o grau de degeneração maior será a pontuação na escala. A Tabela 1 mostra a escala de Pfirrmann..

Porém, por ter como base a inspeção visual da imagem do disco intervertebral, a classificação de Pfirrmann está sujeita a ocorrência de variações intra e interobservador nas classificações feitas em momentos diferentes ou por diferentes profissionais, quando observado o mesmo disco intervertebral (6). Por este motivo, nos últimos anos, muitos pesquisadores têm buscado formas de diagnosticar anormalidades da coluna vertebral e suas estruturas usando métodos quantitativos automáticos ou semi-automáticos. Michopoulou et al. (7) propôs a utilização de descritores baseados em textura para a quantificação da degeneração do disco, e a associação de tais características com a idade do paciente. Alomari et al. (8) desenvolveu um método para a segmentação de hérnia de disco, usando um modelo de vetor de gradiente de contorno ativo e a classificação dos discos com hérnia ou não por meio de um classificador Bayesiano, obtendo um nível médio de acurácia de 92,5%. Oktay et al. (9) realizou um trabalho que detecta se o disco está degenerado ou não, este trabalho analisou os discos intervertebrais de 102 voluntários obtendo uma acurácia de 92,81%. A pesquisa de Jong (10) classificou o nível de degeneração segundo os graus definidos por Pfirrmann (4) de 30 discos intervertebrais, obtendo uma taxa de acerto de 50%.

Tabela 1. Classificação da degeneração do disco intervertebral, adaptado de Pfirrmann et al. 2001 (4).

Grau	Estrutura	Distinção entre núcleo e ânulo	Intensidade do sinal	Altura do disco intervertebral
I	Homogêneo, branco brilhante	Nítida	Branco Brilhante	Normal
II	Não homogêneo, com ou sem bandas horizontais	Clara	Branco Brilhante	Normal
III	Não homogêneo, cinza	Não Clara	Branco para Cinza	Normal à discreta diminuição
IV	Não homogêneo, cinza para preto	Ausente	Cinza para Preto	Normal à moderada diminuição
V	Não homogêneo, preto	Ausente	Cinza para Preto	Espaço discal colapsado

Seguindo essa tendência, o estudo aqui relatado teve por objetivo correlacionar medidas quantitativas de textura e forma extraídas de imagens ponderadas em T2 com as respectivas análises de degeneração discal baseadas na classificação de Pfirrmann.

Material e métodos

Foram utilizadas imagens de RM de 210 discos da região lombar da coluna, extraídos de exames de 98 indivíduos voluntários, adultos, de ambos os sexos, com idade entre 20 e 80 anos e estão divididos em dois grupos. Um grupo possui 85 indivíduos adultos assintomáticos, saudáveis, sedentários, de ambos os sexos, com idade entre 20 e 40 anos. O outro grupo é formado por 13 indivíduos sintomáticos, sedentários, de ambos os sexos, com idade entre 30 e 80 anos. A inclusão do segundo grupo foi necessária para contemplar o grau V de degeneração discal. Com isso a base de imagens ficou composta por 26 discos classificados com grau 1, 67 com grau 2, 66 com grau 3, 30 com grau 4 e 21

com grau 5. Os exames de RM foram realizados em aparelho de Ressonância Magnética de 1.5 Tesla, modelo ACHIEVA da Philips instalado no Serviço de Radiodiagnóstico do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto-USP. Os parâmetros de aquisição foram: plano sagital, FOV = 22x22cm, espessura de corte = 4mm, número de cortes = 16, tempo de echo = 120ms, tempo de repetição = 3900ms, tamanho da imagem = 512x512 pixels, obtendo as imagens ponderadas em T2. As imagens originais em formato DICOM (Digital Imaging and Communications in Medicine) foram convertidas para o padrão JPEG.

Os discos intervertebrais foram classificados por dois médicos radiologistas especialistas de acordo com Pfirrmann (4), nos casos onde houve discordância entre os dois profissionais, um terceiro radiologista com 16 anos de experiência em RM de coluna vertebral definiu a correta avaliação, e esta classificação foi utilizada como padrão "ouro" para a validação do algoritmo computacional desenvolvido.

Para este estudo foi aplicado o algoritmo desenvolvido por Barreiro et al. (11), que utiliza imagens da área central da coluna vertebral ponderada em T2, analisando a condição dos discos localizados nas regiões L1-L2, L2-L3, L3-L4, L4-L5 e L5-S1. Esta análise se inicia com a segmentação semi-automática de cada disco, com a posterior extração dos atributos de textura definidos por Haralick (12), dos momentos invariantes definidos por Hu (13), e extração de outros atributos como altura, largura, e localização na imagem gerando um vetor de características com 39 atributos para cada disco.

Com o objetivo de reduzir a dimensionalidade deste vetor, foi realizada a seleção de atributos, utilizando uma rede neural artificial (RNA) do tipo multilayer perceptron (MLP) como avaliador de subconjunto e busca exaustiva como o método de pesquisa, obtendo-se um vetor com apenas 6 atributos, que são: altura, energia, média da soma, primeiro momento invariante, momento C00 e momento C11. Estes atributos são utilizados como valores de entrada de outra RNA que faz a classificação dos discos.

Para o treinamento da RNA de classificação, os 210 discos foram separados em um conjunto de treinamento e outro de teste, sendo que 2/3 dos dados (140) foram utilizados para treinamento e 1/3 dos dados (70) foram utilizados para teste. Como o conjunto de dados não é homogêneo, foi feita uma separação proporcional das amostras de cada subgrupo segundo os cinco graus da classificação de Pfirrmann, ficando os conjunto de treinamento e de teste de acordo com as figuras 1a e 1b.

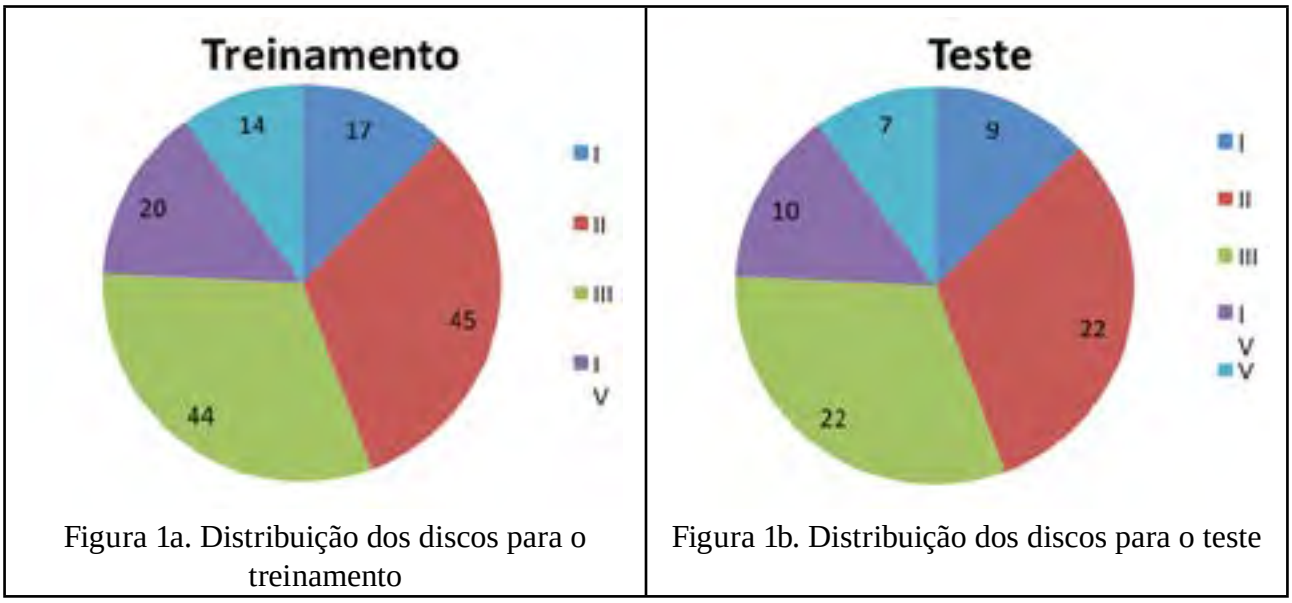


Figura 1a. Distribuição dos discos para o treinamento

Figura 1b. Distribuição dos discos para o teste

A RNA criada para a classificação deste estudo é do tipo MLP, com algoritmo de treinamento do tipo backpropagation. A arquitetura da MLP é definida por 6 neurônios na camada de entrada, 15 neurônios na camada intermediária e 1 neurônio na camada de saída. Essa arquitetura foi selecionada

empiricamente, de forma exaustiva, após sucessivas tentativas para a determinação do melhor desempenho. As 6 entradas da rede estão relacionadas aos seis atributos extraídos das imagens ponderadas em T2.

Após treinar a rede, o conjunto de teste foi submetido à RNA. Cada saída obtida com o conjunto dos seis atributos provenientes de cada disco, foi comparada a classificação de Pfirrmann esperada e o calculado do erro padrão para a proporção de acertos foi feita considerando-se a equação 1.

$$erro = Z * \frac{\alpha}{2} * \sqrt{\frac{\hat{p} * (1 - \hat{p})}{n}} \quad (1)$$

Onde:
Z = Valor tabelado da distribuição normal padronizada,
α = é o nível de significância,
$\hat{p}$ = é a proporção de acerto,
n = tamanho da amostra.

Resultados

Nos experimentos realizados com todos os 39 atributos, a média no nível de acerto foi de 76,19% com erro padrão da proporção de 9,97%. Porém com os experimentos utilizando os 6 atributos indicados no procedimento de seleção de atributos, a média no nível de acerto foi de 81,42%. Considerando o cálculo do intervalo de confiança da proporção pela distribuição normal padronizada, pode-se afirmar com 95% de confiança, que a taxa de acerto está entre 72,31% e 90,53%, sendo o erro padrão da proporção de 9,11%.

A Tabela 2 apresenta os resultados da classificação, as (TP) taxas de verdadeiro-positivos para os diferentes níveis de degeneração variou entre 0,742 e 0,905, com uma média ponderada das taxas de TP de 0,814. As áreas sob a curva ROC (AUC) para os diferentes níveis de degeneração variou entre 0,896 e 1,000, com uma média ponderada de 0,935.

A Tabela 2. resultados da classificação usando os vetores obtidos com os 6 atributos. (TP = True--positivo, FP = Falso-positivo, AUC = Área sob a curva ROC, W. Avg. = Média ponderada)

A Tabela 2. resultados da classificação usando os vetores obtidos com os 6 atributos. (TP = True-positivo, FP = Falso-positivo, AUC = Área sob a curva ROC, W. Avg. = Média ponderada)

	TP rates	FP rates	AUC
Grau I	0,885	0,000	0,993
Grau II	0,821	0,133	0,919
Grau III	0,742	0,111	0,896
Grau IV	0,833	0,022	0,961
Grau V	0,905	0,000	1,000
W. Avg.	0,814	0,080	0,935

Discussão

Os resultados iniciais indicam que o método proposto apresenta bom potencial, com valores de acerto elevados.

A seleção de atributos dos vetores obtidos permitiu uma redução nos dados em quase sete vezes comparado com o número de atributos nos vetores de atributos completo, além de obter um melhor resultado.

As limitações do presente estudo estão na distribuição não uniforme de casos entre os níveis de degeneração do disco e o número total de discos utilizados, apenas 210.

Conclusão

Foi apresentado um método para classificar atributos quantitativos extraídos de imagens ponderadas em T2 com as respectivas classificações de Pfirrmann, em imagens de RM de discos intervertebrais, utilizando-se atributos de textura e forma e redes neurais artificiais. Os resultados indicam que a redução de dimensionalidade teve um efeito positivo sobre os resultados da classificação. As AUC indicam sensibilidade e especificidade satisfatórias dos métodos propostos para o conjunto de dados usados, indicando a viabilidade da abordagem proposta para classificação semi-automática da degeneração do disco intervertebral em imagens de RM ponderadas em T2.

Agradecimentos

Esta pesquisa foi parcialmente financiada pela CAPES, CNPq, FAPESP, IFTM e pelo Conselho de Pesquisa em Engenharia e Ciências Naturais do Canadá.

Referências

[1] Takatalo J, Karppinen J, Taimela S, Niinimäki J, Laitinen J, Blanco Sequeiros R, et al. Body mass index is associated with lumbar disc degeneration in young Finnish males: subsample of Northern Finland birth cohort study 1986. BMC Musculoskelet Disord [Internet]. 2013;14:87. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3599904&tool=pmcentrez&rendertype=abstract

[2] Kerttula L, Kurunlahti M, Jauhiainen J, Koivula A, Oikarinen J, Tervonen O. Apparent diffusion coefficients and T2 relaxation time measurements to evaluate disc degeneration : A quantitative MR study of young patients with previous vertebral fracture. Acta radiol [Internet]. 2001 Jan;42(6):585–91. Available from: http://acr.sagepub.com/lookup/doi/10.1080/028418501127347241

[3] Luoma K, Vehmas T, Riihimäki H, Raininko R. Disc height and signal intensity of the nucleus pulposus on magnetic resonance imaging as indicators of lumbar disc degeneration. Spine (Phila Pa 1976). 2001;26(6):680–6.

[4] Pfirrmann CW, Metzdorf A, Zanetti M, Hodler J, Boos N. Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine (Phila Pa 1976). 2001;26(17):1873–8.

[5] Stelzeneder D, Welsch GH, Kovács BK, Goed S, Paternostro-Sluga T, Vlychou M, et al. Quantitative T2 evaluation at 3.0T compared to morphological grading of the lumbar intervertebral disc: A standardized evaluation approach in patients with low back pain. Vol. 81, European Journal of Radiology. 2012. p. 324–30.

[6] Nagy SA, Juhasz I, Komaromy H, Pozsar K, Zsigmond I, Perlaki G, et al. A Statistical Model for Intervertebral Disc Degeneration: Determination of the Optimal T2 Cut-Off Values. Clinical Neuroradiology. 2013;1–9.

[7] Michopoulou S, Costaridou L, Vlychou M, Speller R, Todd-Pokropek A. Texture-based quantification of lumbar intervertebral disc degeneration from conventional T2-weighted MRI. Acta radiol. 2011;52(1):91–8.
[8] Alomari RS, Corso JJ, Chaudhary V, Dhillon G. Toward a clinical lumbar CAD: Herniation diagnosis. Int J Comput Assist Radiol Surg. 2011;6(1):119–26.
[9] Oktay AB, Albayrak NB, Akgul YS. Computer aided diagnosis of degenerative intervertebral disc diseases from lumbar MR images. Comput Med Imaging Graph. 2014;38(7):613–9.
[10] Jong JJA de. Automatic quantification of intervertebral disc degeneration assessed by T2-weighted magnetic resonance imaging. 2012.
[11] Barreiro MS, Nogueira-Barbosa MH, Rangayyan RM, de Menezes Reis R, Pereyra LC, Azevedo-Marques PM. Semiautomatic classification of intervertebral disc degeneration in magnetic resonance images of the spine. 5th ISSNIP-IEEE Biosignals Biorobotics Conf Biosignals Robot Better Safer Living [Internet]. 2014 May;1–5. Available from: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6880984
[12] Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. IEEE Trans Syst Man Cybern. 1973;3(6).
[13] Hu MK. Visual Pattern Recognition by Moment Invariants. IRE Trans Inf Theory. 1962;8(2):179–87.

CONTATOS

e-mail: marcelobarreiro@iftm.edu.br