

Comunicação

[Communication]

Sistema computacional para auxílio ao diagnóstico em exames de tuberculose animal

[A computer software system to assist in the diagnosis of animal tuberculosis]

A.C.N.R. Gracioso¹, R.A.M. Souza Filho¹, A. Gonzaga², F.J.R. Fernandez³

¹Faculdade de Tecnologia - Carapicuíba, SP

²Universidade de São Paulo - Escola de Engenharia de São Carlos, SP

³Universidade de São Paulo - Escola Politécnica de São Paulo, SP

O Instituto Biológico de São Paulo (IB) desenvolve e realiza pesquisas para o negócio agrícola nas áreas de sanidade animal e vegetal, bem como suas relações com o meio ambiente. Um dos serviços oferecidos é a realização de exames laboratoriais para diagnóstico de doenças na área animal. O Laboratório de Tuberculose do IB é responsável pela pesquisa e diagnóstico de tuberculose em animais. A tuberculose é a principal causa de morte por um único agente, sendo que a tuberculose animal causada pelo bacilo bovino (*Mycobacterium bovis*) é transmissível para humanos, principalmente pela ingestão de alimentos contaminados e pela inalação de aerossóis. Por esse motivo, a detecção de animais infectados é uma preocupação constante do Ministério da Agricultura, Pecuária e Abastecimento (MAPA) (Ruggiero *et al.*, 2007; Benesi *et al.*, 2008).

Atualmente esta detecção depende de análise microscópica, efetuada por um técnico especializado, que busca por padrões da infecção em uma amostra de tecido. Desta forma, o estudo de técnicas que auxiliem o diagnóstico, como as técnicas CBIR (*Content-based Image Retrieval*), constitui um campo promissor por tornar possível executar buscas às bases de imagens existentes, baseadas em critérios de similaridades de forma automatizada.

Em sistemas CBIR, as imagens são descritas por seus atributos nos vetores de características. Esses vetores contêm a representação da imagem (Müller *et al.*, 2004). Características como cor,

forma e textura, extraídas de imagens, sintetizam propriedades inerentes a ela. Essas propriedades são utilizadas no processo de indexação e recuperação do conjunto de imagens (Gonzalez e Woods, 2005). As transformadas *Wavelets* (TW) são mecanismos para decompor ou quebrar sinais nas suas partes constituintes, permitindo analisar dados em diferentes domínios de frequência com a resolução de cada componente associada à sua escala. As imagens naturais, imagens médico-veterinárias, por exemplo, possuem regularidades que podem ser detectadas por meio de uma decomposição *wavelet* (Mallat, 1984; Daubechies, 1992). Neste caso, um número significativo de coeficientes tem pequenas magnitudes, podendo estes ser descartados com pouca distorção na imagem. O objetivo da TW é descorrelacionar os *pixels* de cada subimagem ou compactar o máximo possível à informação em um número menor de coeficientes (Castañón, 2003).

Quanto ao classificador, os sistemas de inferência *fuzzy* permitem o tratamento de informações incertas e imprecisas, as quais estão representadas por uma família de conjuntos *fuzzy*. Esses sistemas de inferência oferecem uma forma sistemática para a modelagem de processos cujas informações são fornecidas de forma qualitativa. Dentro deste contexto, a representação do sistema pode ser realizada por meio de expressões linguísticas, descritas por conjuntos *fuzzy* que expressam o comportamento do processo. O processo de inferência *fuzzy*, ou raciocínio aproximado, permite o mapeamento do conhecimento a respeito de um sistema por

Recebido em 14 de abril de 2011

Aceito em 22 de maio de 2012

E-mail: carol.nrg@gmail.com

meio de regras *fuzzy* do tipo “se-então”. Mediante um conjunto finito dessas regras, pode-se determinar, pelo processo de inferência, o comportamento das variáveis de saída do sistema (Tsoukalas e Uhrig, 1997).

O grande problema dos sistemas baseados em regras está focalizado na quantidade de parâmetros a serem ajustados para um funcionamento adequado do modelo. Esses parâmetros livres se encontram na formação das funções de pertinência dos antecedentes e dos consequentes. Para se criar uma regra nebulosa do tipo “se algo for alto então a saída é baixa”, precisa-se definir que tipo de função e ajustar os parâmetros que indicam o que é “alto” e “baixo” no contexto das regras. Esse procedimento de ajuste não é trivial (Silva *et al.*, 2010).

Os sistemas híbridos neuro-*fuzzy* unem a flexibilidade dos sistemas *fuzzy*, provida pelas regras, com a arquitetura de redes neurais (Silva

et al., 2010) que propicia o aprendizado adaptativo por meio de dados de treinamento, ajustando adaptativamente os parâmetros das funções de pertinência (Braga *et al.*, 2007). Mais especificamente, a arquitetura NEFCLASS (Fig. 1) infere regras e ajusta o perfil das funções de pertinência dos antecedentes (Klose *et al.*, 2000).

Neste trabalho, a TW foi utilizada na extração das características de textura das imagens para compor o vetor de características delas, e um sistema NEFCLASS foi utilizado na classificação. O sistema foi desenvolvido utilizando-se o *software* Matlab, versão 7.0. O banco de imagens foi composto por 141 imagens, sendo 84 com tuberculose e 57 sem tuberculose (Fig. 2 e 3). As imagens foram cedidas pelo Laboratório de Tuberculose do Instituto Biológico de São Paulo. Todas as imagens são provenientes de exames de microscopia de linfonodo e pulmão, possuem tamanho de 1280 x 1024 *pixels* e codificação de 24 *bits*.

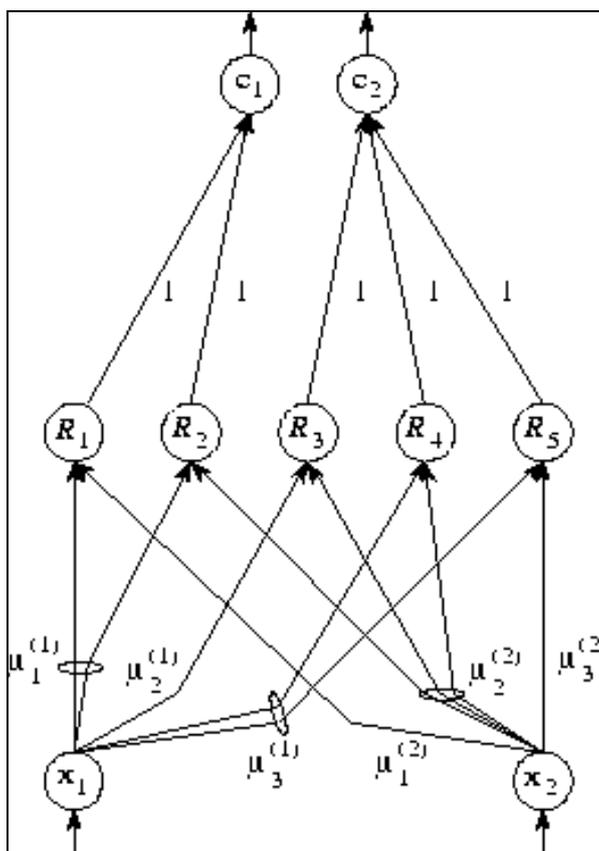


Figura 1. Exemplo NEFCLASS com duas entradas.

Fonte: Klose *et al.*, 2009.

Antes do início do processamento, as imagens foram recortadas, extraindo-se sua parte central com dimensão de 512 x 512, e convertidas para escala de cinza com oito *bits* de profundidade. Para representar as imagens, consideraram-se apenas suas características de textura, obtidas por meio da transformada *Wavelet* de Haar. Uma vez aplicada a transformada *Wavelet* Haar nas imagens,

coeficientes de detalhe e de aproximação do sétimo nível de decomposição foram armazenados nos vetores de características, gerando vetores com 16 valores cada. A apresentação dos dados para o Sistema neuro-fuzzy NEFCLASS foi feita de forma *off-line*, e nas regras foram utilizados pesos fixos com valor igual a um.

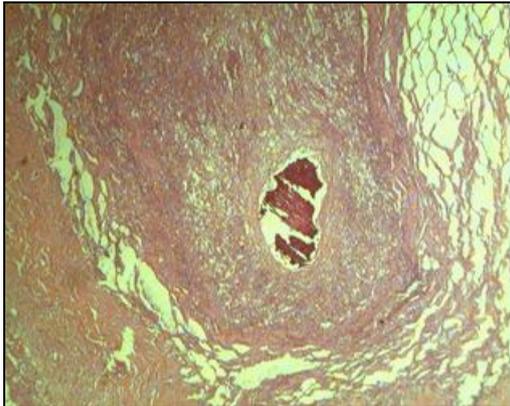


Figura 2. Imagem corte histológico de pulmão bovino positivo para tuberculose corado pelo H.E.

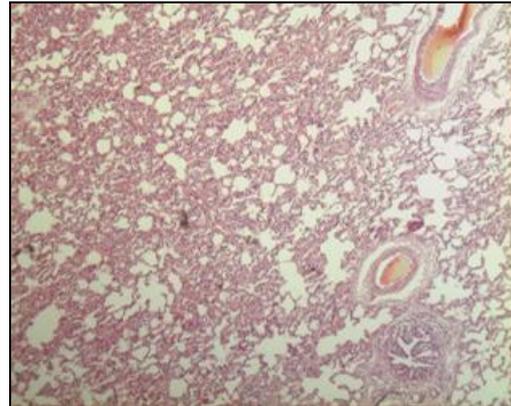


Figura 3. Imagem corte histológico de pulmão bovino negativo para tuberculose corado pelo H.E.

Os testes realizados utilizando-se o método *leaving-one-out* (Stone, 1977) para classificação das imagens entre patológicas e não patológicas para imagens de exames de tuberculose apresentaram uma precisão de recuperação de 72,34%, sendo que, deste percentual, 13,47% das imagens foram classificadas com grau de pertinência igual a um, ou seja, dado $m_A(x) : x \rightarrow [0,1]; x \in X$, em que $m_A(x)$ retorna

o grau de pertinência do elemento x , pertencente ao universo de discurso X , em relação ao conjunto fuzzy A , o grau de pertinência é dado por $m_A(x) \hat{=} [0,1]$, em que os valores limites zero e um indicam, respectivamente, exclusão total e inclusão total no conjunto. A Tab. 1 apresenta os resultados obtidos na classificação.

Tabela 1. Resultado da classificação

	Imagens de tuberculose					
	Positivo (qtd/%)			Negativo (qtd/%)		
	Total	Acerto	Pertinência 1	Total	Acerto	Pertinência 1
qtd	84	63	12	57	39	7
%	100	75	14,28	100	68,42	12,28

A utilização de funções de pertinência de tipos diferentes, como trapezoidal, por exemplo, poderia aumentar o índice de inclusão total, permitindo ao sistema classificar maior quantidade de imagens com segurança, contudo os valores apresentados já representariam uma diminuição da quantidade de exames com o respectivo laudo emitido por especialistas.

Dados os valores obtidos nos testes apresentados na Tab. 1, considerando-se VP um elemento de entrada positivo classificado como positivo, VN um elemento de entrada negativo classificado como negativo, FP um elemento de entrada negativo classificado como positivo e FN um elemento de entrada positivo classificado como negativo, a sensibilidade apresentada pelo classificador, dada por $VP/(VP+FN)$, foi de 75%. A razão de falsos positivos $FP/(VN+FP)$ foi de

32%, e a razão de falsos negativos FN/(VP+FN) foi de 25%. O classificador apresentou uma especificidade VN/(FP+VN) de 68%, precisão VP/(FP+VP) de 77% e acurácia (VP+VN)/(VP+VN+FP+FN) de 72%.

Apesar de a especificidade ser baixa devido ao número elevado de FP (12,76%), a precisão e a acurácia indicam que a metodologia é eficiente, pois os VP também são elevados (44,68%). Portanto, se novos classificadores forem implementados, resultados mais eficientes poderão ser atingidos.

Diante destes resultados, é importante destacar que as imagens não passaram por nenhum pré-processamento que visasse melhorar sua qualidade e que o estudo de técnicas que ressaltem as características mais relevantes para a caracterização da imagem poderia levar a

melhores resultados. As imagens utilizadas nos testes foram obtidas por microscopia óptica; imagens obtidas por meio de microscopia eletrônica tendem a aumentar as taxas de acerto, uma vez que possuem melhores resoluções. Além disso, outras técnicas de extração de características podem ser testadas, como, por exemplo, técnicas que levam em consideração a cor. Neste trabalho, foram consideradas apenas as características de textura das imagens, e a cor é um fator bastante relevante para a classificação das imagens em exames de tuberculose.

Ainda, para melhorar os resultados, apresentam-se passíveis de investigação duas vertentes: a melhoria dos atributos que caracterizam as imagens e o estudo de ajustes no classificador.

Palavras-chave: CBIR, *wavelets*, neuro-fuzzy, tuberculose, diagnóstico

ABSTRACT

The results obtained in evaluating the efficiency of a Neuro-Fuzzy System NEFCLASS (Neuro-Fuzzy Classification) in image classification of cattle tuberculosis, based on its texture features extracted using the wavelet transform are presented. For testing, images of animal tissues diagnosed with tuberculosis were used, as provided by the Tuberculosis Laboratory at the Instituto Biológico de São Paulo. The results of this study can serve as a basis for developing systems for diagnosis aimed at reducing human effort, by automating all or parts of the classification of images, helping lab technicians to sort amongst different pathologies.

Keywords: CBIR, wavelets, neuro-fuzzy, tuberculosis

AGRADECIMENTOS

À Doutora Eliana Roxo, do Laboratório de Tuberculose do Instituto Biológico de São Paulo, pela disponibilização das imagens de tecidos de animais diagnosticados com tuberculose.

REFERÊNCIAS

BENESI, F.J.; PINHEIRO, S.R.; MAIORKA, P.C. *et al.* Relato de caso: Tuberculose em Caprino (*Capra Hircus*). *Arq. Inst. Biol.*, v.75, p.217-220, 2008.

BRAGA, A.P.; CARVALHO, A.P.L.F.; LUDEMIR, T.B. *Redes Neurais Artificiais, Teoria e Aplicação*. LCT, 2007. 238p.

CASTAÑÓN, C.A.B. *Recuperação de Imagens por Conteúdo através de uma Análise Multirresolução por Wavelets*. 2003. 91f. Dissertação (Mestrado). Instituto de Ciências Matemáticas e Computação – ICMC. Universidade de São Paulo, São Carlos, 2003.

DAUBECHIES, I. *Ten Lectures on Wavelets*, vol. 61. CBMS-NSF Regional Conference Series in Applied Mathematics, Rutgers University and AT&T Bell Laboratories, 1992. 377p.

GONZALEZ, R.C.; WOODS, R.E. *Processamento de Imagens Digitais*. São Paulo: Edgard Blücher, 2005.

KLOSE, A.; KRUSE, R.; SCHULZ, K.; THÖNNESEN, U. Controlling Asymmetric Errors in Neuro-Fuzzy Classification. In: Symposium on Applied Computing, SAC .15., 2000, *Proceedings...* New York: ACM, 2000. p.505-509.

MALLAT, S.G. A Theory for Multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pat. Anal. Mach. Intel.*, v.11, p.674-693, 1989.

MÜLLER, H.; MICHOUX, N.; BANDON, D.; GEISSBUHLER, A.A. Review of content-based image retrieval systems in medical applications – clinical benefits and future directions. *Int. J. Med. Informat.*, v.73, p.1-23, 2004.

RUGGIERO, A.P.; IKUNO, A.A.; FERREIRA, V.C.A.; ROXO, E. Tuberculose bovina: alternativas para o diagnóstico. *Arq. Inst. Biol.*, v.74, p.55-65, 2007.

SILVA, I.; SPATTI, D.H.; FLAUZINO, R. *Redes Neurais Artificiais para Engenharia e Ciências Aplicadas*. Artliber, 2010. 399p.

STONE, M. An asymptotic equivalence of choice of model by cross-validation and akaike's criterion. *J. Royal Stat. Soc. Series B (Methodological)*, v.39, p.44-47, 1977.

TSOUKALAS, L.H.; UHRIG, R.E. *Fuzzy and Neural Approaches in Engineering*. v.10. Michigan: wiley, 1997. 587p.