

# Oráculo gráfico como apoio na avaliação de sistemas de auxílio ao diagnóstico

Fátima L. S. Nunes<sup>1</sup>, Márcio E. Delamaro<sup>2</sup>, Rafael A. P. Oliveira<sup>2</sup>

<sup>1</sup>Escola de Artes, Ciências e Humanidades – EACH  
Universidade de São Paulo (USP) São Paulo – SP – Brasil

<sup>2</sup>Instituto de Ciências Matemáticas e de Computação – ICMC/USP  
Universidade de São Paulo (USP) São Paulo – SP – Brasil

fatima.nunes@usp.br, {delamaro,rafael}@icmc.usp.br

***Abstract.** One of the greatest challenges at developing computer aided diagnosis (CAD) systems is to ensure its correction. This paper presents an approach that uses concepts from Content-Based Image Retrieval to construct a practical methodology to evaluate CAD systems.*

***Resumo.** Um dos grandes desafios no desenvolvimento de sistemas de auxílio ao diagnóstico é a garantia de que os resultados estejam corretos. Este artigo apresenta uma abordagem que usa conceitos de Recuperação de Imagens Baseada em Conteúdo para estabelecer uma metodologia pragmática para avaliação de sistemas de auxílio ao diagnóstico.*

## 1. Introdução

Os sistemas de diagnóstico auxiliado por computador (*computer-aided diagnosis* – CAD) consistem em sistemas computacionais com a finalidade de auxiliar na tomada de decisão a respeito de um diagnóstico (Nunes, 2006). O diagnóstico auxiliado por computador é definido por Giger (2000) como aquele no qual o radiologista usa os resultados de uma análise computadorizada de imagens médicas como uma “segunda opinião” na detecção de lesões e na elaboração do diagnóstico.

Esses sistemas constroem uma saída a partir de informações provenientes de diversas fontes, sendo que as imagens médicas de diversas modalidades são as principais. São altamente específicos porque devem considerar as particularidades de cada doença a ser diagnosticada e cada modalidade médica geradora de imagens.

Para atingir patamares aceitáveis e superar esses desafios, tais sistemas precisam ser adequadamente avaliados. No entanto, devido à dificuldade de estabelecer métricas padronizadas, em geral as avaliações são empíricas e poucas abordagens têm sido apresentadas utilizando técnicas de Engenharia de Software.

Neste artigo é apresentada a proposta de uma nova abordagem para auxiliar na avaliação de sistemas CAD, que faz a junção de conceitos de oráculos de teste de software e recuperação de imagens baseada em conteúdo. Nas Seções 2 e 3 são apresentados conceitos que compõem a base teórica do trabalho. A proposta do trabalho é apresentada na Seção 4. Finalmente, a Seção 5 apresenta as considerações finais do artigo, com reflexões sobre o potencial da abordagem apresentada.

## 2. Recuperação de imagens baseada em conteúdo

As buscas tradicionais em bases de imagens utilizam palavras-chaves (Lieberman, Rosenzweig e Singh, 2001) que consistem em atributos textuais, numéricos, ou

similares. Esta forma tradicional torna difícil a pesquisa quando se deseja comparar, por exemplo, uma imagem fornecida pela saída de um programa com outra já conhecida.

O conceito de CBIR (*Content-Based Image Retrieval*) vem suprir esta necessidade. São sistemas computacionais que visam a localizar em uma base de imagens aquelas mais similares a uma imagem de consulta, de acordo com um ou mais critérios fornecidos, obtidos a partir da extração de características da imagem. Dentro do contexto deste artigo, um sistema de CBIR é composto basicamente por três partes: extratores, funções de similaridades e estruturas de indexação (Smeulder et al., 2003).

Os extratores são métodos computacionais que retiram características das imagens a partir de algoritmos que analisam cores, formas, texturas ou outros aspectos relacionados à imagem toda ou a parte dela. O conjunto de características de uma imagem forma o seu vetor de características, utilizado na sua indexação e recuperação. Outro elemento que exerce influência nos resultados da busca é a escolha de medidas de similaridade entre as imagens. Para aplicar uma consulta por similaridade é necessário usar uma função de distância, definindo-se a função de similaridade (Vasconcelos, 2004), um algoritmo com a função de comparar dois vetores de características e retornar um valor não negativo. Quanto menor o valor retornado, maior é a semelhança entre a imagem modelo e a imagem procurada.

### **3. Teste de software e oráculos**

A atividade de teste vem se tornando cada vez mais presente nas organizações que desenvolvem software, dada a crescente necessidade por produtos de qualidade. Por outro lado, estudos (Harrold, 2000) mostram que na virada do século passado 50% do esforço para desenvolvimento eram gastos com a atividade de teste em softwares não críticos. Para softwares críticos essa porcentagem pode ser bem superior. Grande interesse tem sido empregado no desenvolvimento de técnicas, estratégias e ferramentas que permitam um aumento na qualidade e na produtividade desta atividade. Em particular, a automatização, por meio de uso de ferramentas de apoio tem sido de vital importância e concentrado esforços de pesquisadores e desenvolvedores.

Teste de software é uma atividade dinâmica, que consiste em selecionar casos de teste, executá-los com o programa e verificar se o comportamento apresentado é o esperado, de acordo com alguma especificação. Um **oráculo de teste** é um mecanismo capaz de decidir se uma execução está correta ou não, ou seja, se apresenta ou não os resultados esperados. Muitas vezes o papel de oráculo é desempenhado pelo próprio testador que deve consultar a especificação e dar o veredicto.

A automatização de oráculos de teste pode auxiliar os desenvolvedores a alcançarem os objetivos mencionados, provendo qualidade e produtividade, minimizando a necessidade de intervenção humana no processo de teste. Um ponto particularmente delicado na automatização de um oráculo, explorado neste trabalho, aparece quando as saídas a serem comparadas se encontram num formato não trivial, como uma interface gráfica ou imagem produzida pelo programa sendo considerado.

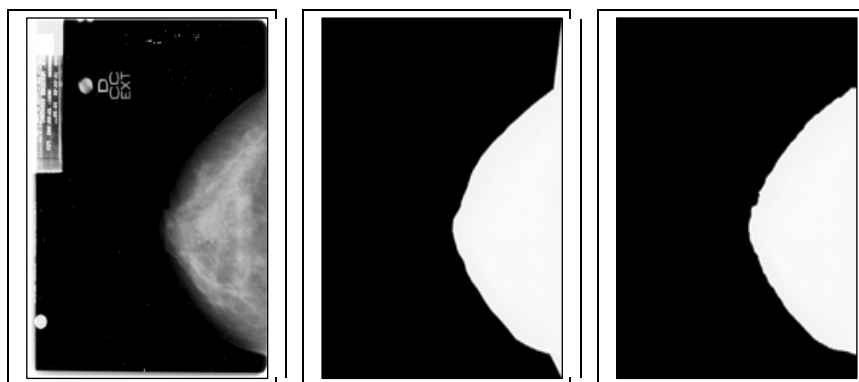
### **4. Avaliação de CAD usando oráculos automatizados**

A composição de um diagnóstico envolve alta dose de subjetividade, considerando fatores como experiência do radiologista e fadiga existente no momento da leitura de imagens médicas. Sistemas CAD podem contribuir para reduzir essa subjetividade, chamando a atenção do especialista para áreas suspeitas.

Neste trabalho propõe-se a automatização de oráculos de programas cuja saída é apresentada na forma de uma imagem, no nosso caso, especificamente, as imagens resultantes de um sistema CAD. Em resumo, tem-se uma ou várias imagens armazenadas que representam a saída esperada de uma execução de um programa CAD e tem-se um segundo conjunto de imagens, resultado real da execução desse programa e das quais se deseja verificar a correção. No nosso caso, o primeiro conjunto é constituído por imagens examinadas por usuários especialistas, geralmente radiologistas com experiência na área de aplicação. Dependendo do caso, essas imagens podem ser avaliadas por mais de um especialista e uma média das avaliações pode ser obtida.

Deve-se notar que a comparação exata não é suficiente para o propósito apresentado, visto que levaria em conta somente a configuração espacial dos pixels presentes em uma imagem. Além das dificuldades de casamento espacial de tais imagens, esta comparação eliminaria possibilidades da flexibilidade necessária quando o fator humano é introduzido no processo. Para viabilizar o proposto foi construído um núcleo de sistema que permite o desenvolvimento de extratores e funções de similaridade sem grande esforço e seu acoplamento em um sistema de CBIR que permite a avaliação desejada (Oliveira, Delamaro e Nunes, 2008).

Atualmente está sendo realizado um estudo de caso no qual dois conjuntos de imagens mamográficas foram manualmente marcados, com a finalidade de identificar a região correspondente à mama. Esses conjuntos são utilizados como referência para comparação com o resultado de um programa que faz essa identificação de modo automático, como parte de um CAD previamente desenvolvido. A Figura 1 mostra uma imagem mamográfica original (a), a mesma imagem após segmentação automatizada para extração das características (b) e após marcação manual, usada como oráculo (c).



**Figura 1. (a) imagem mamográfica original; (b) imagem segmentada automaticamente e (c) imagem marcada manualmente.**

A comparação está sendo realizada utilizando três extratores (área, perímetro e assinatura) e a distância euclidiana como função de similaridade (Vasconcelos, 2004). O extrator de área faz a contagem da quantidade de pixels presentes dentro da borda da mama e o extrator de perímetro contabiliza os pixels que fazem parte da borda demarcada. O terceiro extrator implementado é uma medida de forma e verifica a diferença entre os formatos das bordas demarcadas, também considerando a marcação automatizada e manual. Sua implementação consistiu basicamente em discretizar a borda demarcada, selecionando-se pontos igualmente espaçados, considerando um ângulo pré-determinado formado entre o ponto central da área da mama e a fronteira

identificada. A Tabela 1 mostra os valores obtidos para a distância entre as duas últimas imagens exibidas na Figura 1, considerando cada um dos extratores isoladamente.

**Tabela 1. Valores dos extratores empregados para a imagem da Figura 1.**

Imagem/Extrator	Área	Perímetro	Assinatura
Distância	0,002584	0,001720	0,000292

Com esse estudo de caso pretende-se mostrar a factibilidade desta abordagem, contribuindo para a automatização do processo de avaliação de sistemas CAD ou outros programas que produzam imagens como resultado de sua execução.

## 5. Considerações finais

Conforme apresentado, este artigo propõe uma forma de determinar uma metodologia de avaliação de sistemas CAD, utilizando conceitos de oráculos de teste e CBIR. Os resultados preliminares indicaram a adequação do sistema em desenvolvimento. O estudo de caso aqui apresentado mostra que a utilização do sistema pode ser ampliada, bastando que sejam implementados extratores adequados.

Propondo-se uma forma mais pragmática de avaliação, é possível estabelecer uma metodologia que torne a avaliação de CAD menos susceptível à base de imagens utilizada nos testes e a erros humanos inseridos quando a comparação é realizada com base apenas no sistema de visão humano.

## 6. Referências

- Giger, M. L. (2000) “Computer-aided diagnosis of breast lesions in medical images”, *Computing in Science & Engineering*, v. 2, n. 5, p. 39-45.
- Hafner, J., Sawhney, H., Equitz W., Flickner, M. , Niblack, W. (1995) “Efficient color histogram indexing for quadratic form distance functions”, *IEEE Trans. Pattern. Anal. Machine Intell.*, v. 17, p. 729–736.
- Harrold, M. J. (2000) “Testing: a roadmap”, In *ICSE'00: Proceedings of the Conference on The Future of Software Engineering*, p. 61–72, Nova York, EUA.
- Lieberman, H., Rosenzweig, E. and Singh, P. (2001) “Aria: An Agent for Annotating and Retrieving Images”, *IEEE Computer*, v. 34, n. 7, p. 57–62.
- Nunes, F. L. S. (2006) “Introdução ao Processamento de Imagens Médicas para Auxílio ao Diagnóstico”, In *Atualizações em Informática*, 1<sup>ed</sup>, Rio de Janeiro: PUC-Rio, v.1, p. 73-126.
- Oliveira, R. A. P., Delamaro, M. E. and Nunes, F. L. S. (2008). “Estrutura para Utilização de Recuperação de Imagens Baseada em Conteúdo em Oráculos de Teste de Software com Saída Gráfica”, In *Anais do IV Workshop e Visão Computacional*, v.1, p. 205-210.
- Smeulder, A., Worring, M., Santini, S., Gupta, A. and Jain, R. (2003) “Content based image retrieval at the end of the early years”, *IEEE Trans. Pattern Anal. Mach. Intell.*, v. 22, n. 12, p. 1349–1380.
- Vasconcelos, N. (2004). “On the efficient evaluation of probabilistic similarity functions for image retrieval”, *IEEE Trans. Inf. Theory* v. 50, n. 7, p. 1482–1496.