

# **Estudo Comparativo de Matrizes de Co-ocorrência em Análise de Imagens Médicas: Diferenciação de Tecidos Cólicos**

**Carlos Andres Ferrero<sup>1,3</sup>, Newton Spolaôr<sup>1</sup>, Huei Diana Lee<sup>1</sup>,  
Cláudio Saddy Rodrigues Coy<sup>2</sup>, João José Fagundes<sup>2</sup>, Feng Chung Wu<sup>1,2</sup>**

*<sup>1</sup>Centro de Engenharias e Ciências Exatas – Universidade Estadual do Oeste do Paraná, Laboratório de Bioinformática – LABI, Parque Tecnológico Itaipu – PTI, Foz do Iguaçu, PR, Brasil*

*<sup>2</sup>Faculdade de Ciências Médicas – Universidade Estadual de Campinas, Serviço de Coloproctologia, Campinas, SP, Brasil*

*<sup>3</sup>Instituto de Ciências Matemáticas e de Computação – Universidade de São Paulo, Laboratório de Inteligência Computacional – LABIC, São Carlos, SP, Brasil*

**Resumo** — *O armazenamento de imagens médicas tem incentivado a aplicação de métodos de Análise de Imagens apoiada por Inteligência Computacional para identificar padrões de anormalidades. Este trabalho consiste na aplicação de uma metodologia para diferenciação de tecidos cólicos sobre um conjunto de 134 imagens colonoscópicas. Nesse sentido, é apresentado um estudo comparativo de modelos de classificação de tecidos cólicos construídos a partir de características de textura de imagens baseadas em matrizes de co-ocorrência, considerando as representações de 32, 64, 128 e 256 níveis de cinza. Cada representação foi analisada em relação ao tempo de execução e à precisão preditiva do modelo.*

**Palavras chave** — *inteligência computacional, imagens colonoscópicas, análise de imagens médicas, matriz de co-ocorrência.*

## **1. Introdução**

A utilização de sistemas de gerenciamento de dados nas diversas áreas do conhecimento vem proporcionando o armazenamento de um crescente volume de dados, os quais incluem conteúdo multimídia, como vídeo, imagens e texto [1–3]. O acúmulo de grandes conjuntos de dados tem incentivado a aplicação de métodos computacionais que permitam realizar uma análise mais completa dos dados, no intuito de auxiliar nos processos de tomada de decisões. No contexto de dados do tipo imagem, essa análise pode ser realizada por meio de uma interação entre as áreas de Análise de Imagens – AI – e de Inteligência Computacional – IC –, a qual consiste, primeiramente, na extração de características visuais intrínsecas das imagens e, posteriormente, no reconhecimento dos padrões que permitam classificar novas imagens.

Na área médica, exames são comumente registrados no formato de imagens com o objetivo de complementar os laudos médicos e prover uma descrição mais completa do exame médico e do estado do paciente. Ainda, essas imagens registradas, são utilizadas pelos especialistas para dar suporte ao diagnóstico de enfermidades.

Segundo o Instituto Nacional do Câncer, o câncer colorretal constitui a quarta maior incidência entre todos os tumores malignos, independentemente do sexo [4]. Em relação à mortalidade, as neoplasias do intestino grosso também ocupam o quarto lugar, ocorrendo em menor frequência apenas que tumores de pulmão, estômago e mama [5]. O exame de colonoscopia representa uma ferramenta indispensável para o diagnóstico de doenças do

intestino grosso, sendo assim de interesse a análise de imagens referentes às anormalidades que podem ser evidenciadas por meio desse exame médico.

Este trabalho constitui parte do projeto de Análise de Imagens Médicas, o qual está sendo desenvolvido em uma parceria entre o Laboratório de Bioinformática – LABI – da Universidade Estadual do Oeste do Paraná – UNIOESTE –, o Parque Tecnológico Itaipu – PTI – e o Serviço de Coloproctologia da Faculdade de Ciências Médicas – FCM – da Universidade Estadual de Campinas – UNICAMP.

O projeto de Análise de Imagens Médicas é constituído de duas fases. Na primeira fase, Identificação de Padrões, imagens coletadas a partir de exames médicos, são analisadas e representadas por um conjunto de características, de modo a tornar possível a aplicação de técnicas de Aprendizado de Máquina, uma subárea da IC, que possibilita a busca por padrões em grandes volumes de dados [6]. Na segunda fase, caracterizada como Classificação de Imagens, são construídos sistemas computacionais que permitem classificar novas imagens utilizando os padrões encontrados. Dessa maneira, é possível auxiliar aos especialistas da área nos processos de tomada de decisões associados à identificação de anormalidades presentes em imagens médicas.

Neste trabalho é apresentado um estudo de utilização de atributos de textura baseados em matrizes de co-ocorrência para a construção de modelos de classificação. No trabalho são utilizados conjuntos de dados construídos a partir da extração de características de textura, considerando quatro representações diferentes de Níveis de Cinza – NC –, as quais são comumente utilizadas na literatura para a representação de imagens médicas. Nesse sentido, o objetivo do trabalho consiste em avaliar os modelos construídos, para encontrar a representação de NC que se apresenta mais adequada para representar as imagens e, conseqüentemente, induzir modelos. Na Seção 2 é descrita a metodologia utilizada para a realização dos experimentos. Posteriormente é mostrado na Seção 3 um estudo de caso utilizando um conjunto de imagens de exames de colonoscopia, na qual se compara e discute a qualidade dos modelos construídos e, por fim, são apresentadas as conclusões e os trabalhos futuros na Seção 4.

## 2. Metodologia

A metodologia proposta tem como objetivo diferenciar tecidos cólicos presentes em imagens de exames de colonoscopia, por meio da construção de modelos a partir de informações contidas nas imagens. Essa tarefa é realizada mediante a aplicação de quatro etapas: (1) coleta do conjunto de imagens, (2) construção do vetor de características, (3) construção de modelos para classificação e (4) avaliação de modelos, as quais são descritas a seguir.

### *Etapa 1 – Coleta do Conjunto de Imagens*

Os sistemas de gerenciamento de imagens médicas vêm se tornando comuns em ambientes médicos, permitindo o armazenamento de imagens que possibilitam complementar os diagnósticos realizados pelos especialistas. No caso de exames de colonoscopia, as imagens registram anormalidades do tipo polipóides, diverticulares e vasculares, dentre outras. Desse modo, torna-se interessante oferecer suporte à diferenciação entre tecidos cólicos com e sem presença de anormalidade. Na Etapa 1 da metodologia, a partir de um conjunto de  $n$  imagens, são selecionados fragmentos de imagem que apresentam tecidos cólicos. Para cada imagem são selecionados dois fragmentos: um de anormalidade morfológica e o outro

sem anormalidade. Desse modo, os  $2n$  fragmentos obtidos constituem o Conjunto de Imagens, definido pela notação  $CI=\{Im_1, \dots, Im_{2n}\}$ . A Figura 1 ilustra a seleção de fragmentos de anormalidade e sem anormalidade.

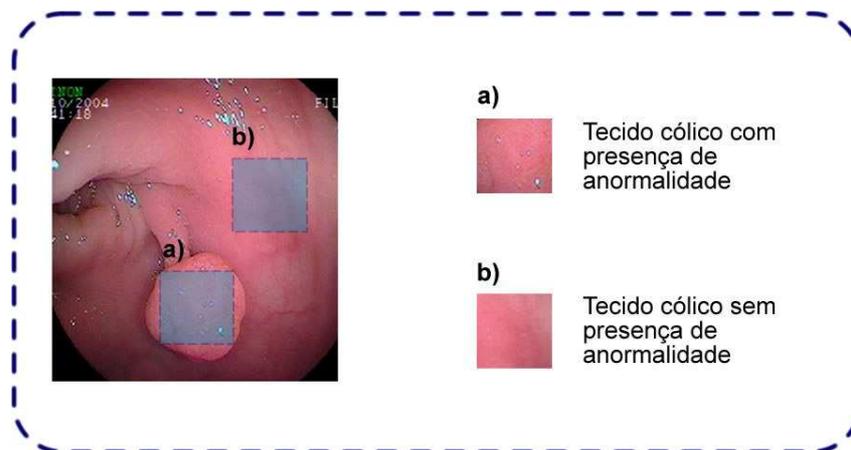


Figura 1. (a) Tecido cólico de anormalidade e (b) tecido cólico sem anormalidade.

#### Etapa 2 – Construção do Vetor de Características

Na análise de imagens é freqüente a descrição de imagens conforme propriedades baseadas em cor, textura e forma [7]. Na Etapa 2, a partir das propriedades de cada imagem pertencente ao conjunto  $CI$ , são extraídas diversas características. Essas informações podem ser representadas no formato atributo-valor, o qual é um padrão freqüente dos dados de entrada de algoritmos para a construção de modelos. As  $d$  características a serem extraídas em cada imagem constituem o vetor de atributos  $V=\{C_1, \dots, C_d\}$ . A  $i$ -ésima imagem processada origina o exemplo  $E_i$ , o qual é descrito por  $E_i=\{A_{i,1}, \dots, A_{i,d}, A_{i,d+1}\}$ , em que os atributos de  $A_1$  a  $A_d$  correspondem aos valores das características definidas em  $V$  e o atributo  $A_{d+1}$  contém a classe do exemplo. A união de todos os exemplos define o Conjunto de Dados Inicial, que é denotado por  $CDI=\{E_1, \dots, E_{2n}\}$ .

Algumas das maneiras mais comuns de representação de imagens médicas são as baseadas em textura, as quais podem ser extraídas por meio da construção de Matrizes de Co-ocorrência – MC – [2,8–10] e de espectros de textura [1,11,12], entre outras técnicas. A aplicação dessas técnicas permite extrair atributos que descrevem as propriedades de disposição e espaço entre níveis de cinza e outros elementos que compõem uma imagem. A extração de características de textura partir de MC, foco deste trabalho, possibilita a representação computacional de imagens de colonoscopia para que seja possível realizar o restante de metodologia. Essas matrizes são constituídas de  $lc$  linhas e colunas, em que  $lc$  indica o número de NC considerado. As linhas e as representam os NC presentes na imagem, de modo que o valor na posição  $(i,j)$  indica o número de transições de nível de cinza  $i$  para nível de cinza  $j$ . Na Figura 2 é apresentado um exemplo de construção de uma MC a partir de uma imagem.

Imagem				
0	0	0	1	2
1	1	0	1	1
2	2	1	0	0
1	1	0	2	0
1	0	1	0	0

		Matriz de co-ocorrência		
		0	1	2
0		8	8	2
1		8	6	2
2		2	2	2

Figura 2. Imagem representada em três NC como uma matriz de pixels 5x5 e a respectiva MC para a direção  $0^\circ$  e a distância 1 [1].

A construção da matriz de co-ocorrência observada na Figura 2 pode ser exemplificada pela obtenção do valor na posição  $(1,0)$ . O valor 8 corresponde à contabilização dos pixels de valor 1, cujo vizinho direito ou esquerdo apresenta valor 0. A observação dos vizinhos direito e esquerdo é indicada pela direção  $0^\circ$  e a distância 1. Desse modo, a partir das MC construídas, são extraídos um conjunto de valores calculados (descritores). Haralick *et al.* (1973) propuseram 11 descritores para representar propriedades de textura das imagens: Segundo Momento Angular, Contraste, Correlação, Variância, Entropia, Momento Inverso da Diferença, Soma da Média, Soma da Variância, Soma da Entropia, Diferença da Variância e Diferença da Entropia [8].

#### Etapa 3 – Construção de Modelos para Classificação

Após a construção do conjunto CDI, é realizada a indução de modelos utilizando algoritmos de Inteligência Computacional. A geração de modelos com base em grandes conjuntos de dados possibilita a classificação de novos exemplos, a partir do aprendizado obtido na resolução bem-sucedida de problemas anteriores. Desse modo, os modelos construídos podem ser úteis para auxiliar especialistas no processo de tomada de decisões [6,13].

#### Etapa 4 – Avaliação de Modelos

Os modelos construídos devem ser avaliados quanto à sua qualidade preditiva, considerando a eficiência que apresentam na predição de novos exemplos. Essa avaliação pode ser realizada, por exemplo, utilizando o método de validação cruzada [14] e a tabela de contingência, entre outras técnicas. A tabela de contingência permite definir as características preditivas de um modelo por meio de quatro métricas [15]:

- Valor Preditivo Positivo – VPP – porcentagem de exemplos de tecidos com presença de pólipos em relação ao total de exemplos classificados como tecidos de pólipos;
- Valor Preditivo Negativo – VPN – porcentagem de exemplos de tecidos sem presença de pólipos em relação ao total de exemplos classificados como tecidos sem pólipos;
- Sensibilidade – porcentagem de exemplos classificados como tecidos com presença de pólipos em relação ao total de exemplos de tecidos de pólipos;
- Especificidade – porcentagem de exemplos classificados como tecidos sem presença de pólipos em relação ao total de exemplos de tecidos sem pólipos.

Outra técnica envolvida nessa etapa corresponde à avaliação e à validação dos modelos por parte de especialistas do domínio. A contribuição dos especialistas auxilia na eliminação de padrões desnecessários e na certificação dos demais padrões, o que proporciona maior robustez preditiva aos modelos construídos. Nessa avaliação podem ser

propostas melhorias nas etapas anteriores, de modo a extrair características de maior significância e, conseqüentemente, formular hipóteses mais consistentes.

As Etapas 1 e 2 são realizadas com o suporte de um sistema computacional desenvolvido no Laboratório de Bioinformática – LABI –, denominado *Medical Image Analysis System* – MIAS [16,17]. A Etapa 3 é desenvolvida com o auxílio da ferramenta WEKA<sup>1</sup>.

### 3. Estudo de Caso

A metodologia apresentada foi aplicada a um conjunto de 67 imagens provenientes de exames de colonoscopia coletadas pelo Serviço de Coloproctologia da Faculdade de Ciências Médicas da UNICAMP. As imagens utilizadas correspondem a exames de colonoscopia nos quais foram identificados pólipos de anormalidade morfológica do Tipo Ip (classificado como Pólipo Protruso Pediculado, conforme as normas da Sociedade Japonesa de Pesquisa do Câncer Colorretal) [4,5]. Inicialmente, para cada imagem, foram selecionados, manualmente, fragmentos de tecido de mucosa cólica de anormalidade e sem presença de anormalidade morfológica (Figura 1). Cada tecido cólico foi selecionado considerando o tamanho da região interna de cada pólipo, de modo a utilizar a maior quantidade de informação possível a respeito da anormalidade morfológica. A partir dessa seleção foi criado o conjunto CI contendo 134 fragmentos de imagens.

Posteriormente, foi realizada a representação das imagens por meio da transformação de cada imagem do formato original (RGB<sup>2</sup>) para o formato em escala de cinza, o qual é adequado para a definição do valor de intensidade de cada ponto (pixel) e para a conseqüente construção da MC. Em relação à representação das imagens em escala de cinza, foram consideradas as representações de 32, 64, 128 e 256 níveis de cinza, no intuito de avaliar qual dessas representações apresenta-se mais adequada na extração de características de textura para a construção de modelos de classificação. Nesse sentido, foi utilizado o aplicativo MIAS para construir as MC correspondentes de cada imagem, considerando uma distância de 1 e as direções 0, 45, 90 e 135 graus. A partir das MC, foram extraídas as 11 características de textura mencionadas na Etapa 2 da metodologia utilizada. Para cada característica, apenas o valor da média dos valores das quatro direções foi considerado. Com isso, foi possível uma representação atributo-valor do conjunto de fragmentos de imagem para cada representação em NC. Na Figura 3 é apresentado um gráfico de barras referente aos tempos de construção das MC e de extração de características das 134 imagens para cada representação em nível de cinza<sup>3</sup>. Nesse gráfico, o eixo das abscissas indica as representações de níveis de cinza e o eixo das ordenadas indica o tempo de execução, medido em segundos, utilizando a escala logarítmica.

---

1 <http://www.cs.waikato.ac.nz/ml/weka/>.

2 Vermelho, verde e azul, do inglês *red, green, blue*.

3 O computador utilizado para a realização dos experimentos possui um processador Pentium IV de 2.8 Ghz, memória de 1.0 GB, com Windows XP *Service Pack 2* e máquina virtual Java JDK 1.6.0.04.

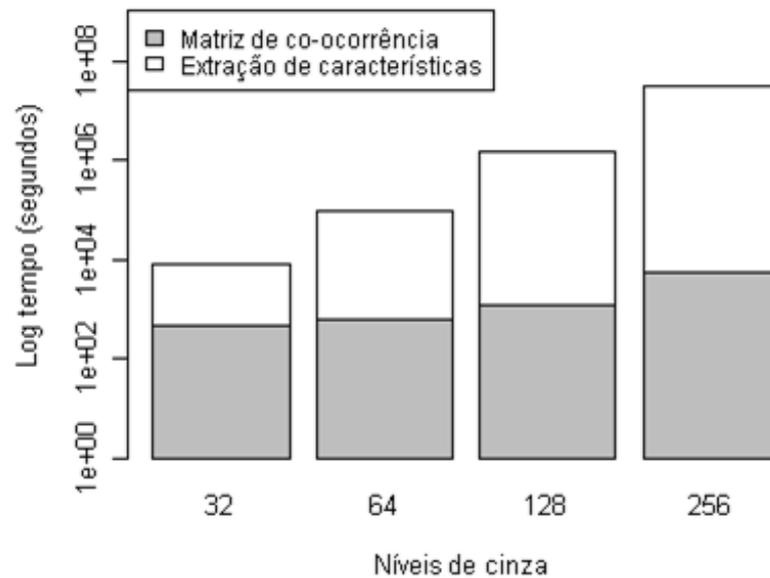


Figura 3. Tempos de construção das MC e de extração de características das 134 imagens para cada representação em NC.

O gráfico apresentado permite evidenciar o crescimento exponencial do tempo de execução em relação ao aumento do número de níveis de cinza, conforme era esperado. A representação em 32 NC apresentou o menor tempo consumido tanto na construção da MC, quanto na extração de características, enquanto que a representação em 256 NC apresentou o maior tempo para ambas as tarefas.

A partir dos dados no formato atributo-valor foram induzidos modelos de classificação. Para cada conjunto de dados foi construída uma árvore de decisão, utilizando o algoritmo J48, uma implementação do algoritmo C4.5 [18] disponibilizada na ferramenta WEKA. Após, os modelos construídos foram avaliados, primeiramente com base na precisão preditiva e, posteriormente, em relação às medidas extraídas das tabelas de contingência. A avaliação em relação à precisão preditiva foi realizada utilizando a técnica de validação cruzada, considerando dez partições. Desse modo, é possível obter uma aproximação do erro verdadeiro dos modelos. Os valores de erro médio e desvio-padrão dos modelos de cada uma das representações de NC são apresentados na Tabela 1.

Tabela 1. Valores de erro médio e desvio padrão dos modelos construídos.

	32 NC	64 NC	128 NC	256 NC
<b>Erro médio</b>	0,2973	0,2258	0,2387	0,2313
<b>Desvio-padrão</b>	0,1401	0,1029	0,0975	0,1091

Conforme os valores apresentados na Tabela 1 pode ser constatado que o menor erro médio foi o obtido pela representação em 64 níveis de cinza e o menor desvio-padrão pela representação em 128 NC, indicando que esse último teve a menor variação da taxa de erro. Porém, a superioridade da representação em 64 NC pode não ser significativa em relação aos outros modelos construídos. Para constatar a diferença entre a qualidade dos modelos foi aplicado um teste estatístico. Visto que as taxas de erro dos modelos passaram pelo teste da normalidade, o teste estatístico *t* para dados não emparelhados foi utilizado para constatar a existência de diferença estatisticamente significativa entre o erro do modelo de

64 NC e os erros dos outros três modelos, considerando um nível de significância de 95%. Os *p*-valores resultantes da aplicação do teste em relação a 32, 128 e 256 NC foram 0,1902, 0,7996 e 0,9046, respectivamente, indicando que não foi possível verificar diferença estatisticamente significativa entre o erro médio do modelo com 64 NC e outros três.

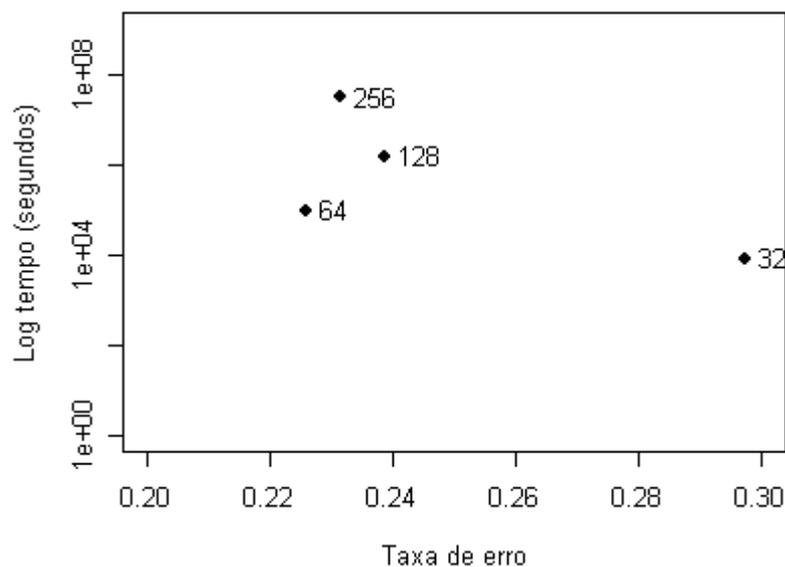


Figura 4. Tempo de execução e erro médio do modelo para cada representação em NC.

A partir desses valores de precisão é importante ressaltar a relação entre o erro médio de cada modelo e o respectivo tempo de execução até a extração de características, apresentado anteriormente. A Figura 4 ilustra o gráfico que relaciona as duas informações. Nesse gráfico, o eixo das abscissas indica a taxa de erro e o eixo das ordenadas indica o tempo de execução das tarefas, em segundos, utilizando a escala logarítmica. Cada um dos quatro pontos do gráfico é identificado, ao lado, pelo número de NC correspondente. Com base nesse gráfico, pode ser observado que a representação em 64 NC originou o modelo de menor erro médio, embora tenha alcançado o segundo menor tempo de execução. Por outro lado o modelo construído a partir da representação em 32 NC apresentou o menor tempo de execução e o maior erro médio.

Posteriormente, foi realizada a análise considerando as medidas de valor preditivo positivo, valor preditivo negativo, sensibilidade e especificidade, que podem ser extraídas a partir das tabelas de contingência de cada um dos modelos construídos (Anexo I). Na Tabela 2 são apresentadas as medidas calculadas a partir das tabelas de contingência dos modelos construídos considerando as representações de 32, 64, 128 e 256 NC.

Tabela 2. Medidas calculadas a partir da tabela de contingência de cada modelo construído.

Medida (%)	32 NC	64 NC	128 NC	256 NC
<b>VPP</b>	70,77	78,46	76,92	80,00*
<b>VPN</b>	69,56	76,81*	75,36	74,32
<b>Sensibilidade</b>	68,66	76,12*	74,63	71,64
<b>Especificidade</b>	71,64	79,10	77,61	82,09*

\* Maior valor de precisão para cada medida.

Em relação ao VPP, pode ser constatado que o modelo construído sobre a representação de 256 níveis de cinza atingiu o valor de 80,00%, sendo o maior valor de probabilidade de classificação correta de imagens de tecidos de pólipos. A especificidade desse modelo, de 82,09%, também superior em relação às especificidades dos demais modelos, indicou maior probabilidade de identificação de imagens sem pólipos. A análise sobre o VPN possibilitou observar que o modelo referente à representação em 64 NC apresentou a maior taxa de acerto, 76,81%, em relação às imagens classificadas como tecidos sem pólipos. Esse modelo apresentou ainda a maior sensibilidade, atingindo 76,12% de probabilidade de identificação de imagens de tecidos de pólipos.

Os modelos obtidos a partir das MC de 128 e 32 NC obtiveram, respectivamente, a terceira e a quarta melhor probabilidade em cada uma das medidas consideradas. De modo geral, os modelos que consideram 64 e 256 NC apresentaram melhores resultados em relação às quatro medidas avaliadas, o que permite afirmar que, embora os valores de VPP e especificidade sejam maiores para o último, o modelo originado da MC de 64 NC apresentou melhor desempenho de VPP e especificidade em relação aos modelos que consideram 32 e 128 NC, indicando o segundo maior valor para ambas as medidas.

#### **4. Conclusão**

De acordo com os resultados apresentados, embora o modelo de representação em 64 NC tenha apresentado o menor erro médio de classificação, não foi possível constatar diferença estatisticamente significativa entre esse e os erros médios de classificação dos modelos construídos a partir de 32, 128 e 256 NC. Por meio da tabela de contingência foi possível verificar o maior poder preditivo do modelo que considera 256 NC, no que se refere às probabilidades de acerto na classificação de tecidos de pólipos e de identificar como tecido sem anormalidade uma imagem de tecido sem pólipos. O modelo gerado sobre a representação de 64 NC apresentou-se como o de maior probabilidade de acerto na classificação de imagens de tecido sem pólipos, e o de maior probabilidade de dada uma imagem de pólipos classificá-la como tal. Ao se observar o tempo de execução para a construção das MC e para a extração das características, é possível afirmar que a representação de 64 NC apresentou menor custo computacional do que a abordagem com 256 NC. Desse modo, a representação em 64 NC apresentou a melhor relação entre o custo computacional e a precisão do modelo correspondente.

Trabalhos futuros incluem a realização de experimentos para a verificação da melhor distância a ser considerada na construção das matrizes de co-ocorrência; a aplicação de métodos de seleção de atributos para procurar subconjuntos de características relevantes e/ou não redundantes; a implementação e utilização de outras características para descrever as imagens de exames de colonoscopia; e a coleta de um número maior de imagens com tipos de anormalidades bem definidas, no intuito de construir modelos melhor adaptados a problemas reais.

#### **Agradecimentos**

Ao Programa de Desenvolvimento Tecnológico Avançado – PDTA/FPTI-BR – pelo auxílio por meio da linha de financiamento de bolsas. À Fundação de Apoio ao Desenvolvimento Científico e Tecnológico do Paraná – ARAUCARIA – pelo auxílio por meio da linha de financiamento de projetos de pesquisa.

## Referências

- [1] KARKANIS, S.; GALOUSSI, K.; MAROULIS, D., “Classification of endoscopic images based on texture spectrum”, *Proceedings of Workshop on Machine Learning in Medical Applications*, 1999, 63–69.
- [2] FELIPE, J. C.; TRAINA, A. J. M.; TRAINA, C., “Retrieval by content of medical images using texture for tissue identification”, *Proceedings of 16th IEEE Symposium on Computer-Based Medical Systems (CBMS 2003)*, 2003, 26–27.
- [3] FERRERO, C. A.; MALETZKE, A. G.; LEE, H. D.; WU, F. C.; NETO, A. P.; FAGUNDES, J. J.; GÓES, J. R. N., “Endia – uma ferramenta para a construção de repositórios de dados de endoscopia digestiva alta”, *X Congresso Brasileiro de Informática em Saúde (CBIS 2006)*, 2006, 1–6.
- [4] QUILICI, F. A., “Colonoscopia”, Lemos Editorial, 2000.
- [5] COTRAN, R. S.; KUMAR, V.; COLLINS, T., “Patologia estrutural e funcional”, Guanabara Koogan, 2000.
- [6] REZENDE, S. O., “Sistemas Inteligentes: Fundamentos e Aplicações”, Manole, 2003.
- [7] HOWARTH, P.; YAVLINSKY, A.; HEESCH, D.; RÜGER, S., “Medical image retrieval using texture, locality and colour”, *Lecture Notes from the Cross Language Evaluation Forum*, 2005, 740–749.
- [8] HARALICK, R.; SHANMUGAM, K.; DINSTEN, I., “Texture features for image classification”, *IEEE Transactions on Systems, Man, and Cybernetics (SMC 1973)*, 3, 1973, 610–621.
- [9] MARQUES, P. M. D. A.; HONDA, M. H.; RODRIGUES, J. A. H.; TRAINA, A. J. M.; TRAINA, C., “Recuperação de imagens baseada em conteúdo: uso de atributos de textura para caracterização de microcalcificações mamográficas”, *Revista Brasileira de Radiologia*, 35, 2002, 93–98.
- [10] KRISHNAN, S. M.; YAP, C. J.; ASARI, K. V.; GOH, P. M. Y., “Neural network based approaches for the classification of colonoscopic images”, *Proceedings of 20th International Conference on IEEE Engineering in Medicine and Biology Society*, 1998, 1678–1680.
- [11] HE, D. C.; WANG, L., “Texture unit, texture spectrum, and texture analysis”, *Proceedings of IEEE Trans. on Geoscience and Remote Sensing*, 28, 1990, 509–512.
- [12] FERRERO, C. A.; LEE, H. D.; CHERMAN, E. A.; COY, C. S. R.; FAGUNDES, J. J.; GÓES, J. R. N.; WU, F. C., “Utilização do espectro de textura para a diferenciação de imagens cólicas em exames colonoscópicos”, *36º Jornadas Argentinas de Informática (36 JAIHO)*, 2007, 1–12.
- [13] WITTEN, H.; FRANK, E., “Data mining: practical machine learning tools and techniques”, Elsevier, 2005.
- [14] APLPAYDIN, E., “Introduction to machine learning”, MIT Press, 2004.
- [15] DORIA, U., “Introdução à bioestatística: para simples mortais”, Elsevier, 1999.
- [16] FERRERO, C. A.; LEE, H. D.; CHERMAN, E. A.; COY, C. S. R.; FAGUNDES, J. J.; GÓES, J. R. N.; WU, F. C., “Utilização de atributos baseados em textura para a caracterização de tecidos cólicos em imagens de colonoscopia”, *VII Workshop de Informática Médica (WIM 2007)*, 2007, 1–10.
- [17] FERRERO, C. A.; LEE, H. D.; WU, F. C.; COY, C. S. R.; FAGUNDES, J. J.; GÓES, J. R. N., “Identificação de padrões de anormalidades em tecidos cólicos”, *I Congresso da Academia Trinacional de Ciências*, 2006, 1–1.
- [18] QUINLAN, J. R., “C4.5: programs for machine learning”, Morgan Kaufmann, 1993.

### **Dados de Contato**

*Carlos Andres Ferrero. Laboratório de Inteligência Computacional, Universidade de São Paulo. Av. Trabalhador São-carlense, 400, 13560-970, São Carlos, SP, Brasil. anfer86@gmail.com.*

*Newton Spolaôr. Laboratório de Bioinformática, Universidade Estadual do Oeste do Paraná. Av. Presidente Tancredo Neves, 6731, 85866-900, Foz do Iguaçu, PR, Brasil. newtonspolaor@gmail.com.*

*Huei Diana Lee. Laboratório de Bioinformática, Universidade Estadual do Oeste do Paraná. Av. Presidente Tancredo Neves, 6731, 85866-900, Foz do Iguaçu, PR, Brasil. hueidianalee@gmail.com.*

*Cláudio Saddy Rodrigues Coy. Serviço de Coloproctologia da Faculdade de Ciências Médicas, Universidade Estadual de Campinas. Rua Carlos Chagas, 420, 13031-970, Campinas, SP, Brasil. ccoy@terra.com.br.*

*João José Fagundes. Serviço de Coloproctologia da Faculdade de Ciências Médicas, Universidade Estadual de Campinas. Rua Carlos Chagas, 420, 13031-970, Campinas, SP, Brasil. jffagundes@mpcnet.com.br.*

*Feng Chung Wu, Laboratório de Bioinformática, Universidade Estadual do Oeste do Paraná. Fundação Parque Tecnológico Itaipu. Av. Presidente Tancredo Neves, 6731, 85866-900, Foz do Iguaçu, PR, Brasil. wufc@pti.org.br.*