

# Uma Metodologia para Auxiliar no Processo de Mapeamento de Formulários Médicos para Bases de Dados Estruturadas

André Gustavo Maletzke<sup>1</sup>, Huei Diana Lee<sup>2</sup>, Wu Feng Chung<sup>3</sup>, Edson Takashi Matsubara<sup>4</sup>,  
Cláudio Sady Rodrigues Coy<sup>5</sup>, João José Fagundes<sup>6</sup>, Juvenal Ricardo Navarro Góes<sup>7</sup>

<sup>1,2,3</sup>Laboratório de Bioinformática — LABI,  
Universidade Estadual do Oeste do Paraná — UNIOESTE, Brasil  
<sup>3,5,6,7</sup>Serviço de Coloproctologia da Faculdade de Ciências Médicas — FCM,  
Universidade Estadual de Campinas — UNICAMP, Brasil  
<sup>2,3,4</sup>Laboratório de Inteligência Computacional — LABIC,  
Universidade de São Paulo — USP, Brasil

**Resumo** - A informação no atual cenário de desenvolvimento tecnológico é fundamental para a tomada de decisões e a capacidade de obtê-la torna-se essencial. O processo de Descoberta de Conhecimento em Bases de Dados pode auxiliar na análise e na compreensão de dados armazenados em uma base de dados. Nesse contexto, hospitais e clínicas médicas utilizam, freqüentemente, documentos em papel para registrar informações sobre pacientes. Para que essas informações possam ser usadas por esse processo é necessário que sejam mapeadas para formatos apropriados, como o atributo-valor, o qual pode ser obtido a partir de bases de dados estruturadas. Neste trabalho é apresentada uma metodologia e a implementação de um protótipo de sistema para auxiliar o mapeamento de informações contidas em formulários médicos impressos relacionados à doença de Crohn, para bases de dados estruturadas.

**Palavras-chave:** Metodologia, Descoberta de Conhecimento em Bases de Dados, Doença de Crohn.

**Abstract** - In the scenario of technological advance, information plays an important role in the decision making process and the capability to obtain it becomes essential. The Knowledge Discovery from Databases process can give support to the analysis and the comprehension of the data contained in databases. In this context, it is very frequent in hospitals and medical clinics the use of printed documents to register information about patients. To turn possible the application of this information to such process, it is necessary to map the data to more appropriate formats, like the attribute-value one, which can be obtained from structured databases. In this work, it is presented a methodology and the implementation of a prototype system to help the mapping of information contained in printed medical forms, related to the Crohn disease, to the structures databases.

**Key-words:** Methodology, Knowledge Discovery from Databases, Crohn Disease.

## Introdução

O avanço tecnológico ocorrido nos últimos anos tem motivado o desenvolvimento e o estudo de diversos métodos e processos para apoiar a análise de dados com o auxílio de ferramentas computacionais [1]. Isso se deve à grande quantidade de dados que vêm sendo armazenada, em uma velocidade cada vez maior, tornando a análise desses dados uma tarefa crescentemente complexa por meio de ferramentas usuais. Um dos processos de apoio à essa tarefa é o processo de Descoberta de Conhecimento em Bases de Dados — DCBD — [2], o qual possibilita a extração de padrões existentes nos dados, por meio por exemplo de métodos de Inteligência Artificial — IA —, que poderão ser utilizados em processos de tomada de decisão.

Diversas áreas podem ser beneficiadas com a aplicação de processos como o DCBD, tal qual a área de medicina. Porém, nessa área específica, há uma grande quantidade de dados que não estão disponíveis em formatos apropriados para a aplicação desse processo. Esses dados podem estar representados em laudos médicos ou em formulários impressos, contendo informações sobre, por exemplo, exames ou sintomatologia de pacientes. Para que esses dados possam ser analisados com o auxílio desses métodos, é necessário que sejam mapeados para formatos apropriados, como a representação em Bases de Dados — BD — estruturadas. São diversos os fatores que podem estar associados à não disponibilidade desses dados em BD, como a não existência de computadores em todos os ambulatórios médicos, a consideração por muitos de que a utilização de for-

mulários manuscritos no relacionamento com o paciente o torna menos impessoal ou a necessidade por se manter registros impressos.

Nesse contexto, um dos temas que tem despertado interesse entre os pesquisadores da área médica está relacionado à doenças inflamatórias intestinais, como a doença de Crohn. Essa enfermidade representa um desafio para médicos e pesquisadores, principalmente por apresentar altos índices de morbidade e mortalidade, etiologia indefinida e ausência de cura até os dias atuais [3] [4]. Este trabalho está inserido dentro do projeto de Análise Inteligente de Dados aplicado ao Mapeamento de Dados Médicos — AIDMD — [5], e tem como objetivo apresentar uma metodologia que auxilia no mapeamento semi-automático, para BD estruturadas, de informações descritas em formulários de múltipla escolha preenchidos manualmente, relacionados à enfermidade de Crohn. Este trabalho apresenta também a implementação de um protótipo de sistema da metodologia proposta e uma avaliação sobre a escolha do método de interpolação mais apropriado para o problema apresentado.

## Materiais e Método

O desenvolvimento tecnológico atrelado a inúmeras pesquisas na área de IA possibilitou o desenvolvimento de métodos que podem ser utilizados para a análise e a extração de padrões, a partir de conjuntos de dados. Nesse contexto, foi desenvolvida uma metodologia para auxiliar na coleta de dados por meio do armazenamento, de maneira estruturada em uma BD, a partir de formulários médicos impressos. Esses padrões poderão posteriormente ser analisados utilizando métodos de IA.

Os formulários aplicados estão estruturados na forma de perguntas e respostas. Cada pergunta representa um atributo na BD e as respectivas alternativas, representadas por uma marca, os possíveis valores que serão atribuídos a esse atributo — Figura 1.

1 - Nível de Leucócitos		
Normal <input type="checkbox"/>	Anormal <input type="checkbox"/>	Não apresenta <input type="checkbox"/>
2 - O paciente apresenta diarreia?		
Sim <input type="checkbox"/>	Não <input type="checkbox"/>	

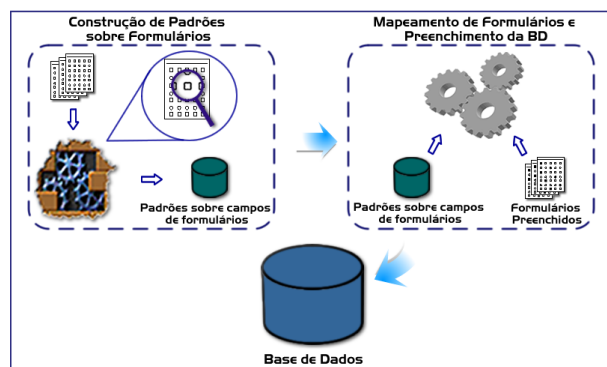
**Figura 1:** Exemplo das perguntas e respostas de um formulário.

Para que esses formulários possam ser tratados computacionalmente, devem ser digitalizados por meio de um digitalizador óptico (*scanner*). Este processo transforma um formulário em um arquivo constituído por elementos de imagem, denominados *pixels*.

A metodologia proposta foi dividida em duas etapas:

1. Etapa 1: Construção de padrões sobre os formulários;
2. Etapa 2: Mapeamento dos formulários e preenchimento da BD.

Cada uma das etapas estão representadas na Figura 2, sendo o objetivo da primeira etapa extrair padrões a partir de formulários modelo. Esses padrões são referentes ao tamanho, formato e localização das marcas dentro de cada formulário ainda não preenchido. Com isso, cada formulário modelo terá seu próprio conjunto de padrões. Este conjunto de padrões será utilizado na segunda etapa, na qual é realizado o mapeamento dos formulários e preenchimento da BD, considerando formulários já preenchidos, por meio da verificação de qual alternativa de cada pergunta foi marcada.



**Figura 2:** Representação esquemática das etapas da metodologia proposta.

Antes de cada uma dessas etapas é necessário que os formulários passem por uma fase de preparação, denominada de pré-processamento.

## Pré-Processamento

Esta fase tem como objetivo preparar os formulários para as etapas de construção de padrões e de mapeamento dos formulários e preenchimento da BD. Para auxiliar nesta fase são aplicados dois métodos de processamento de imagens amplamente utilizados: a binarização e a segmentação [6].

A binarização da imagem de um formulário transforma essa imagem em outra, cujos valores de *pixels* são 0 ou 1. Já o processo de segmentação, consiste em dividir uma imagem tornando-a em outra menor ou em grupo de imagens menores, de maneira que as partes de interesse sejam selecionadas. Desse modo, após o processo de binarização, é realizada a leitura da marca de referência. Essa

marca de referência é representada por uma linha horizontal, situada no cabeçalho de cada formulário e estendendo-se por toda a largura do formulário, e tem como objetivo auxiliar nas operações de correção de possíveis distorções e na segmentação da área de interesse. Para efetuar essas correções, relacionadas à inclinação ou à escala da imagem, são aplicadas operações de rotação e escala, respectivamente. Para ambas as operações, é aplicado um método de interpolação. O processo de segmentação tem como objetivo gerar uma nova imagem a partir da imagem original. A nova imagem é formada pela área retangular, delimitada pelas coordenadas  $(x_0, y_0, x_f, y_f)$ , sendo  $x_0$  e  $y_0$  referentes às coordenadas iniciais da marca de referência, e  $x_f$  e  $y_f$  referentes aos comprimentos da marca de referência e do formulário, respectivamente. Com o término desta fase, o formulário estará no formato apropriado para iniciar-se a primeira etapa da metodologia proposta.

### Etapa 1 : Construção de padrões sobre os formulários

Esta etapa tem como objetivo mapear a localização das áreas de interesse a partir de um formulário modelo (não preenchido) e, armazenar essas informações, de maneira que possam ser utilizadas na etapa de mapeamento dos formulários e preenchimento da BD. O processo de mapeamento das áreas de interesse baseia-se no algoritmo descrito em [7], no qual a idéia básica é explorar o fato de que a grande parte dos documentos digitalizados, possui uma estrutura vertical e horizontal. Esta etapa é dividida em três fases:

1. Fase de segmentação horizontal;
2. Fase de segmentação vertical;
3. Fase de construção de padrões.

#### Fase de segmentação horizontal

Nesta fase, é contabilizada para cada linha horizontal do formulário a quantidade de *pixels* pretos presentes nessa linha, considerando a espessura de uma linha equivalente a um *pixel*. Desse modo, é gerado o Histograma de projeção Horizontal — HH — do formulário, que é representado por um vetor, no qual cada posição armazena a quantidade de *pixels* pretos de cada linha, obtendo-se um vetor de tamanho equivalente à quantidade de linhas existentes. A partir da análise do HH, o formulário é segmentado, por meio da determinação de um limiar, denominado de limiar de segmentação (normalmente 0 ou 1), em banda branca e em banda de texto. Uma banda

branca é constituída por linhas contínuas, nas quais o número de *pixels* pretos é menor ou igual ao limiar estabelecido, e a banda de texto é formada por linhas contínuas contendo uma quantidade de *pixels* pretos maior que o limiar, como ilustrado na Figura 3.

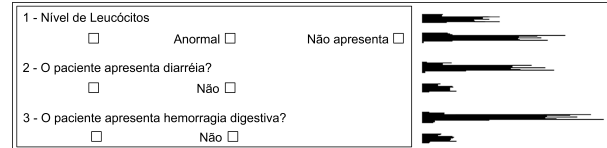


Figura 3: Histograma de projeção Horizontal do formulário.

#### Fase de segmentação vertical

Na fase de segmentação vertical, somente as bandas de texto serão utilizadas, portanto as bandas brancas, nesta fase, são descartadas. Para cada banda de texto é gerado um Histograma de projeção Vertical — HV —, o qual é um vetor de números inteiros, em que cada posição do vetor armazena a quantidade de *pixels* pretos existentes entre as coordenadas  $y$  inicial e final da banda de texto analisada — Figura 4.

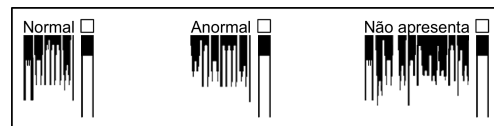


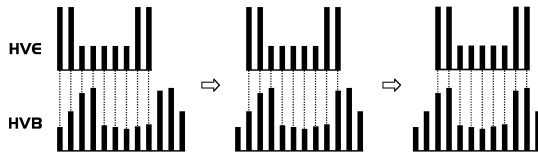
Figura 4: Histograma de projeção Vertical de uma banda de texto.

Cada HV gerado, é particionado, por meio do limiar de segmentação, em um conjunto de HVs denominado de Conjunto de Busca — CB —, no qual cada elemento, denominado HVB, representa um possível objeto do formulário, como marcas e caracteres. Neste trabalho, as marcas de interesse do formulário são representadas pela figura geométrica de um quadrado. Para cada HVB gerado são mantidas suas respectivas coordenadas, as quais serão necessárias na terceira fase.

Ainda nessa fase, é realizada a construção de um conjunto de exemplos. Um exemplo é representado pelo Histograma de projeção Vertical de uma marca — HVE —, que é gerado por meio de uma interface gráfica, na qual deverão ser selecionadas as marcas que irão compor o conjunto de exemplos. Desse modo, para cada marca selecionada é gerado um HVE, sendo este armazenado juntamente com suas respectivas coordenadas, na Base de Exemplos — BE. Esta base será utilizada no processo de reconhecimento das marcas restantes, o qual é realizado na fase de construção de padrões.

## Fase de construção de padrões

Nesta fase é construída uma Base de Padrões — BP —, na qual todas as marcas são mapeadas e suas respectivas coordenadas armazenadas, pois serão necessárias para o preenchimento da BD (Etapa 2 do processo de mapeamento dos dados de um formulário para uma BD). Para isso, todos os HVBs gerados na fase anterior serão submetidos a um classificador, que por meio da BE irá identificar quais HVBs apresentam o padrão de uma marca. Com a finalidade de realizar esse processo foi implementado um algoritmo baseado no método de classificação *Nearest Neighbor* [8]. O objetivo desse método é classificar um exemplo não rotulado por meio da atribuição da classe de seu vizinho mais próximo, contido no Conjunto de Treinamento — CT —, neste caso a BE, gerada na fase anterior. Esse algoritmo compara cada elemento do CB com os elementos contidos na BE. Para realizar essa comparação é utilizado um método baseado na técnica *Dynamic Time Warping* — DTW — [9] que consiste na aproximação de séries que possuem desenvolvimentos semelhantes, porém defasados um em relação ao outro.



**Figura 5:** Técnica utilizada para a comparação de histogramas.

Essa técnica de comparação é apresentada na Figura 5, na qual o histograma superior representa um exemplo da Base de Exemplos (HVE) e o histograma inferior um HVB presente no Conjunto de Busca. A partir dessa técnica, ambos histogramas são comparados e caso sejam semelhantes a comparação é finalizada. Caso contrário, a comparação é realizada novamente, porém entre as frequências  $H_x$  do histograma exemplo e as frequências  $H_{x+t}$  do histograma HVB, com  $x$  variando de 0 até o comprimento do histograma exemplo,  $t$  variando de 1 até  $n$  em intervalos de uma unidade e  $n$  sendo um valor inteiro menor ou igual à diferença de tamanho existente entre o HVE e o HVB.

Para determinar a similaridade entre dois objetos é necessário estabelecer uma medida de distância, que indique quão similares são os objetos em comparação. Neste trabalho foi utilizada a distância  $\chi^2$  [10] para determinar o grau de similaridade entre dois histogramas HVE e HVB, pertencentes a BE e ao CB, respectivamente. Essa distância é dada pela

Equação (1).

$$\chi^2 = \sum_{i=1}^n \left( \frac{Hx[i] - Hy[i]}{Hx[i] + Hy[i]} \right)^2 \quad (1)$$

onde  $Hx(i)$  e  $Hy(i)$  representam cada bin (frequência) normalizado dos histogramas HVE e HVB respectivamente. Com isso, os histogramas do CB que foram semelhantes, de acordo com o DTW usando o  $\chi^2$ , com algum exemplo da BE, terão suas coordenadas mapeadas para a Base de Padrões.

## Etapa 2 : Mapeamento dos formulários preenchidos e preenchimento da BD

A segunda etapa dessa metodologia, tem como objetivo identificar qual alternativa de cada pergunta foi marcada e preencher a BD. O processo de verificação de preenchimento das marcas é constituído por três passos:

1. Realização da segmentação de cada marca do formulário por meio das coordenadas contidas na BP;
2. Contabilização da porcentagem de *pixels* pretos que ocupam a área segmentada;
3. Consideração da alternativa marcada caso a porcentagem seja maior que a porcentagem limiar aceita.

Os formulários são construídos a partir de um Sistema Gerador de Formulários — SGF —, o qual possui uma interface gráfica na qual são fornecidas as perguntas e as alternativas que o formulário deverá conter. Este sistema também é responsável pela construção da BD e do Arquivo de Interpretação do preenchimento do formulário. Este arquivo indica o valor que deverá ser armazenado na BD de acordo com a resposta marcada no formulário. Desse modo, a partir da identificação de uma resposta marcada, a BD é preenchida por meio da análise do arquivo de padrões do formulário. A partir dessa metodologia foi construído um protótipo de sistema. Esse protótipo foi desenvolvido na linguagem de programação *Java*, que utiliza o conceito de Orientação a Objeto e possui ferramentas para auxiliar na manipulação de imagens, como a API *Java Advanced Imaging* — JAI. A persistência dos dados é realizada por meio da linguagem *Structured Query Language* — SQL — e o Sistema Gerenciador de Banco de Dados — SGBD — *mysql* está sendo utilizado para o armazenamento das informações mapeadas a partir dos formulários.

Para determinar o método de interpolação a ser utilizado pelo protótipo do sistema, foram realizados experimentos considerando três tipos de interpolação diferentes presentes na literatura [6]:

- **Nearest Neighbor** — NN: caracteriza-se por não criar novos valores de *pixels*. Este método avalia os valores dos *pixels* vizinhos e atribui o valor de um dos *pixels* vizinhos ao novo *pixel* gerado;
- **Bilinear**: o valor de um novo *pixel* é determinado pelo cálculo da média dos valores de seus vizinhos mais próximos, considerando um conjunto de  $2 \times 2$  vizinhos;
- **Bi-cúbica**: este método cria um novo valor para o *pixel* interpolado atribuindo o valor médio obtido por meio da análise dos valores dos seus vizinhos, considerando um conjunto  $4 \times 4$  vizinhos.

Para a realização dos experimentos foram gerados dez formulários, não preenchidos, com ângulo de inclinação de -5 a -1 graus e de 1 a 5 graus, variando em intervalos de um grau. Foram consideradas imperfeições relacionadas somente à inclinação, a qual foi corrigida aplicando os métodos de interpolação apresentados anteriormente. Para determinar o método de interpolação mais adequado para o problema corrente, foi realizado para cada formulário o seguinte processo:

1. Segmentação das regiões referentes às marcas dos formulários, de acordo com as coordenadas da Base de Padrões;
2. Contabilização para cada região segmentada da percentagem que a borda da marca ocupa em relação à área segmentada.

Foi utilizado como caso de controle um formulário com ângulo de inclinação igual a zero. Neste caso, não foi necessária a realização da operação de rotação, conseqüentemente nenhum método de interpolação foi aplicado. Para a análise dos resultados foi utilizado o teste estatístico não paramétrico e não pareado *Kruskall Wallis* com pós-teste *Dunn*, considerando 95% de significância.

## Resultados

O resultado dos experimentos é apresentado na Tabela 1, na qual são apresentados a média da percentagem de ocupação da borda em relação à área segmentada com seu respectivo desvio padrão para cada variação na angulação do formulário referente à aplicação dos diferentes métodos de interpolação. Na última linha, é apresentada a média e o desvio padrão para cada método de interpolação. Para o caso de controle a média de *pixels* pretos por área segmentada foi igual a 35,68% com desvio padrão de 4%.

Tabela 1: Resultados do experimento.

Formulário	NN	Bilinear	Bi-cúbica
	% Média (DP)	% Média (DP)	% Média (DP)
1	33,28 (5,36)	41,11 (6,93)	39,18 (7,54)
2	36,80 (3,08)	41,85 (4,81)	40,51 (4,75)
3	33,03 (4,17)	41,34 (4,84)	41,19 (4,19)
4	29,38 (2,76)	34,69 (3,83)	37,36 (3,51)
5	35,50 (2,30)	36,88 (3,20)	34,81 (3,07)
6	35,68 (4,00)	35,68 (4,00)	35,68 (4,00)
7	31,86 (3,90)	37,18 (7,15)	40,30 (6,74)
8	34,49 (3,15)	39,05 (6,03)	37,37 (5,90)
9	31,59 (3,55)	36,05 (3,93)	34,50 (3,77)
10	36,76 (2,31)	42,19 (3,06)	40,25 (3,06)
% Média	33,87 (4,12)	39,61 (5,81)	38,96 (5,51)

## Discussão e Conclusões

Durante o processo de digitalização dos formulários algumas imperfeições poderão ocorrer, como alterações na escala e formulários digitalizados com um determinado grau de inclinação. Para reduzir essas imperfeições são utilizados métodos e técnicas de processamento de imagens, como rotação e transformações de escala. A aplicação desses métodos implica na utilização de algum tipo de interpolação, que irá influenciar na qualidade da imagem que está sendo manipulada. Diversos métodos de interpolação são propostos na literatura, sendo que cada método apresenta-se de maneira mais adequada a um determinado domínio ou aplicação. Desse modo, neste trabalho foram realizados experimentos com métodos de interpolação amplamente citados na literatura, com o objetivo de avaliar qual método apresenta-se mais adequado para este problema.

A análise dos resultados demonstrou que não há diferença significativa entre o método de interpolação NN e o caso de controle ( $p > 0,05$ ). Comparações entre os métodos Bilinear e Bi-cúbico também não apresentaram diferença significativa ( $p > 0,05$ ). Para as demais comparações a diferença foi significativa ( $p < 0,001$ ). Portanto, foi determinado o método de interpolação NN como método a ser utilizado pelo protótipo do sistema, pois apresentou-se estatisticamente semelhante ao caso de controle. Esse resultado está relacionado à característica desse método, no qual não são criados novos valores de *pixels*. Desse modo, este método obtém resultados melhores para imagens que apresentam bordas quadradas e bem definidas. Porém, para imagens fotográficas, este método não é adequado, pois esses tipos de imagens possuem transições gradativas entre determinadas áreas, produzindo uma pequena assimetria nas bordas presentes na imagem. O método NN é amplamente aplicado no processamento de imagens médicas, pois dependendo da análise que será realizada sobre a imagem, a geração de novos valores

de *pixels* irá influenciar diretamente na precisão dos resultados.

A partir da aplicação dessa metodologia na construção de um protótipo de sistema, será possível realizar o mapeamento de formulários cuja estrutura e organização das informações sejam desconhecidas. Tornando o processo de mapeamento uma tarefa menos custosa e mais eficiente, se comparada com métodos em que todas as regiões de interesse deverão ser manualmente mapeadas.

O protótipo de sistema desenvolvido possibilitará o preenchimento de uma base de dados estruturada, que posteriormente será submetida a métodos de Inteligência Artificial, associados ao processo de DCBD, para que possíveis padrões presentes nos dados contidos nessa BD sejam encontrados.

Neste trabalho foi apresentada uma metodologia para a construção de um protótipo de sistema, que auxilia no mapeamento de dados contidos em formulários médicos para uma BD estruturada. Essa metodologia será aplicada sobre formulários relacionados às doenças inflamatórias intestinais, especificamente a doença de Crohn. Essa enfermidade tem despertado amplo interesse por parte de pesquisadores, pois pouco sabe-se sobre sua etiologia e fisiopatologia, além de afetar um número crescente de pessoas em todo mundo. Desse modo, a partir da base dados gerada, poderão ser aplicados algoritmos de IA para a extração de padrões contidos nos dados mapeados, com o objetivo de auxiliar médicos e pesquisadores nas pesquisas relacionadas a essa doença. Outras áreas do domínio médico poderão aplicar esta metodologia para mapear informações descritas em formulários para bases de dados estruturadas.

Como trabalho futuro propõem-se a incorporação a esta metodologia, de um módulo para reconhecimento de caracteres manuscritos, o qual possibilitará maior flexibilidade no preenchimento dos formulários e enriquecerá ainda mais os padrões gerados a partir da BD construída por esta metodologia.

## Agradecimentos

Ao Programa de Desenvolvimento Tecnológico Avançado — PDTA/FPTI-BR — pelo auxílio por meio da linha de financiamento de bolsas.

## Referências

- [1] Rezende, S.O. (2003), *Sistemas Inteligentes: Fundamentos e Aplicações*, Barueri — SP, Brasil: Manole.

- [2] Fayyad, U., Piatetsky-Shapiro, G., Smyth, P. (1996), "From data mining to knowledge discovery in databases", *AI Magazine*, p. 37-54.
- [3] Cordeiro, F., Filho, J.S.M., Prolla, J.C. (1994) *Endoscopia Digestiva*, Rio de Janeiro — RJ, Brasil: MEDSI.
- [4] Contran, R., Kumar, V., Collins, T. (2000) *Patologia Estrutural e Funcional*, Rio de Janeiro — RJ, Brasil: Guanabara Koogan.
- [5] Honorato, D.F., Lee, H.D., Monard, M.C., Wu, F.C., Machado, R.B., Ferrero, C.A. (2005), "Uma metodologia para auxiliar no processo de construção de base de dados estruturadas a partir de laudos médicos", In: *V Encontro Internacional de Inteligência Artificial*, Porto Alegre — RS, Brasil, p. 593-601.
- [6] Gonzalez, R.C., Woods, R.E. (2002) *Digital image processing*, Upper Saddle River — NY, USA: Prentice Hall.
- [7] Minh, T.H., Bunke, H. (1992) *Analysis and Understanding of GIRO Check Forms*, Relatório Técnico, Universidade de Berna, Suíça, 37p.
- [8] Mitchell, T.M. (1997) *Machine learning*, Boston, USA: McGraw-Hill.
- [9] Felipe, J.C., Traina, A.J.M. (2004), "Warp Metric Distance: Aprimorando o Uso de Histogramas de Intensidade para Busca por Similaridade em Imagens Médicas", In: *IX Congresso Brasileiro de Informática em Saúde*, Ribeirão Preto — SP, Brasil, v. 1, p. 1-6.
- [10] Zhang, H., Malik, J. (2003) "Learning a discriminative classifier using shape context distance", In: *Proc. of the IEEE CVPR*, Madison — WI, USA: v. 1, p. 242-247.

## Contato

A. G. Maletzke — andregustavom@gmail.com  
H. D. Lee — huei@unioeste.br  
F. C. Wu — wufc@unioeste.br  
E. T. Matsubara — edsontm@icmc.usp.br  
C. S. R. Coy — ccoy@terra.com.br  
J. J. Fagundes — jfagundes@mpcnet.com.br  
J. R. N. Góes — rgoes@mpcnet.com.br

Laboratório de Bioinformática — LABI, Universidade Estadual do Oeste do Paraná — UNIOESTE. Av. Tarquínio Joslin dos Santos, 1300, Pólo Universitário, CEP 85870-650, Foz do Iguçu — PR.