

# Reconhecimento e Recuperação de Imagens

## Utilizando Redes Neurais Artificiais do Tipo MLP

Túlio Cléber Bender<sup>1</sup>, Fernando Santos Osório<sup>1</sup>

<sup>1</sup>Programa Interdisciplinar de Pós-Graduação em Computação Aplicada – Universidade do Vale do Rio dos Sinos (UNISINOS)  
Avenida Unisinos, 950 – São Leopoldo – RS – Brasil

tulio@euler.unisinos.br, osorio@exatas.unisinos.br

**Abstract.** *This paper presents an approach to the task of classification and retrieval of bitmap images using artificial neural networks and color as primary attribute. We also discuss about the impact of separations surfaces drawn by the classifier during training and retrieval and we present a new way to handle this situation.*

**Resumo.** *Este artigo apresenta uma abordagem para reconhecimento e recuperação de imagens estáticas do tipo bitmap baseado em atributos de cor utilizando redes neurais artificiais do tipo Multilayer Perceptron. Também é discutido como as superfícies de separação do classificador determinadas pelos pesos dos neurônios influenciam na precisão do reconhecimento e apresentamos novos mecanismos de como abordar este problema.*

### 1. Introdução

Atualmente com o aumento considerável de bases de dados de imagem e vídeo devido às inovações tecnológicas tornou-se necessário manter algum tipo de mecanismo de indexação para posterior recuperação de dados. Executar esta indexação manualmente é uma tarefa sensível a interpretações subjetivas, é laboriosa e sujeita a erros [ANTANI 02, SEBE 01, BRUNELLI 00, RUI 97, SMITH 96]. Métodos automáticos de indexação e recuperação baseados em algum tipo de característica (cor, forma, textura ou região de uma imagem) são interessantes neste contexto porque reduzem a intervenção humana. A idéia de recuperação de imagens por exemplos é uma extensão da idéia de recuperação de páginas web que contenham no texto as palavras-chave especificadas na pesquisa (por exemplo, Google, Altavista e outros). No caso de imagens, busca-se recuperar em uma base de dados de imagens aquelas que contenham o elemento gráfico procurado como, por exemplo, uma determinada cor, textura, forma ou a combinação destes elementos.

Definimos *imagem-exemplo* como sendo um exemplo que contém o elemento gráfico procurado durante uma recuperação e que pode ser uma imagem completa, região de uma imagem, cor, forma ou textura. As regiões de uma imagem que não são o alvo da recuperação são denominadas de *imagem-de-fundo*. As *características* extraídas de uma imagem são atributos derivados da cor, textura e ou da forma e que denotam alguma propriedade da imagem que pode ser utilizada para comparação entre imagens. A *comparação de características* é uma função que retorna o grau de diferença entre as características extraídas das imagens.

As técnicas de recuperação de imagens vem sendo desenvolvidas desde 1970 principalmente pelas áreas de gerenciamento de banco de dados (indexação textual) e visão computacional [RUI 97]. No método baseado em indexação textual existe um DBMS que gerencia a informação semântica anexada por uma pessoa a imagem. Já nos Content Based Image Retrieval (CBIR) [ANTANI 02] procura-se extrair automaticamente das imagens características que as identifiquem. A partir destas características da imagem-exemplo procura-se por características similares nas imagens disponíveis em uma base de pesquisa. As melhores similaridades são apresentadas como resultado da pesquisa. Um exemplo de um sistema CBIR é o IKONA [IKONA 02], que está disponível na Internet para demonstrações.

## 2. Abordagem Usual Para Recuperação de Imagens

A abordagem utilizada nos sistemas CBIR pode ser dividida em dois subsistemas: geração da base de dados de imagens de pesquisa e o subsistema de pesquisa nesta base. No subsistema de geração da base de dados são aplicados os métodos de extração de características nas imagens e seu resultado é armazenado nesta base. O subsistema de pesquisa aplica os mesmos métodos de extração de características na imagem-exemplo e depois executa uma comparação entre as características extraídas da imagem-exemplo e das armazenadas na base de dados. A comparação das características está baseada numa função que retorna a diferença ou distância entre as características extraídas da imagem-exemplo com as características de cada imagem da base de procura, portanto **executando uma comparação 1-a-1**. As imagens com as menores diferenças são então retornadas ao usuário da pesquisa como sendo as melhores candidatas. A escolha dos métodos de extração e comparação utilizados, que melhor se adaptem a uma determinada base de dados, é normalmente uma responsabilidade de quem projeta e prepara a base de imagens a ser pesquisada.

## 3. Aplicação da Inteligência Artificial

Neste trabalho aplicamos técnicas de Inteligência Artificial (IA) para executar o reconhecimento e recuperação de imagens bitmap a partir de características extraídas automaticamente das imagens. A aplicação faz uso de técnicas de *Machine Learning* da IA com aprendizado supervisionado, onde se optou pelo uso de Redes Neurais Artificiais (RNA) do tipo Multilayer Perceptron (MLP) com algoritmo de aprendizado construtivo e incremental do tipo *Cascade-Correlation* [FAHLMAN 91].

A escolha desta técnica é interessante pela capacidade das RNAs em obter uma generalização de um padrão aprendido a partir de uma base de exemplos. Através do uso de RNAs se pretende obter uma generalização do(s) modelo(s) de imagem procurado e para isto foi desenvolvido o programa *ImageMatching* (IMatch) que implementa esta abordagem. Este programa se baseia principalmente no atributo da cor para executar a tarefa de reconhecimento, pois a identificação baseada nas cores apresenta propriedades interessantes como a independência da posição do padrão procurado na imagem, relativa independência entre a escala do padrão e o fundo da imagem e da orientação do padrão na imagem.

Podemos citar como aplicação do programa IMatch seu emprego nas áreas de visão computacional e robótica, levantamento automático da ocorrência de logomarcas e

localização de objetos em uma cena. A seguir iremos descrever como este programa opera e a metodologia empregada.

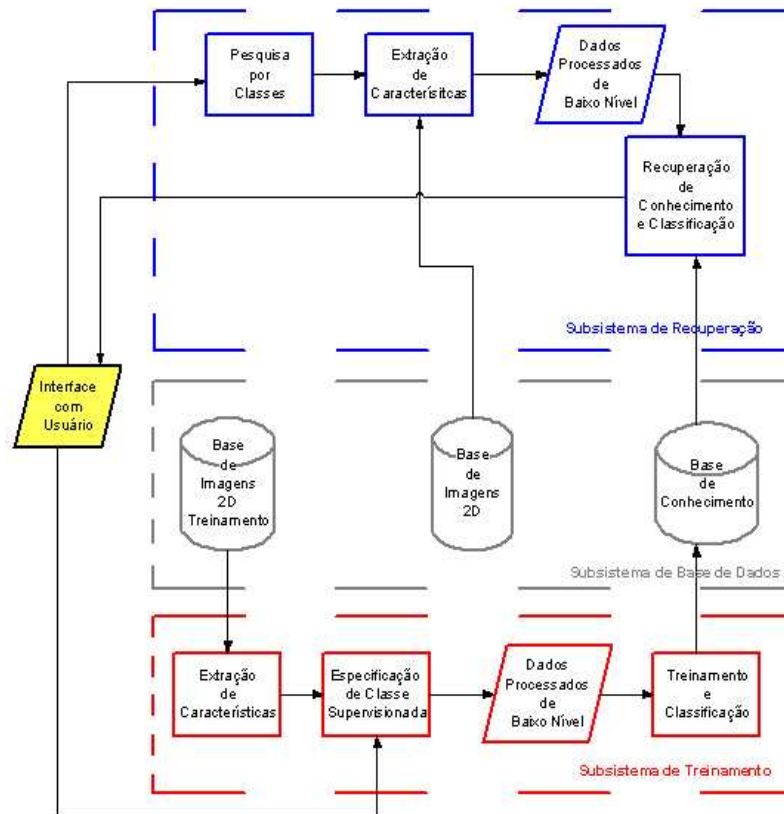
### 3.1. Método de reconhecimento utilizando redes neurais artificiais

O programa IMatch realiza a extração de características das imagens utilizando a cor como atributo básico (a versão atual do sistema não considera a forma dos objetos presentes em uma cena). Sistemas unicamente baseados na cor irão recuperar imagens com uma composição similar de cor, independentemente do contexto das imagens serem diferentes. Apesar desta limitação, trabalhar com as cores possui algumas vantagens em relação à extração de características baseadas em textura ou forma. A cor é uma das características mais utilizadas nos sistemas CBIR porque é relativamente independente quanto ao tamanho da região de interesse em relação à imagem-de-fundo, rotação, escala, e oclusão parcial da imagem-exemplo, sendo computacionalmente menos cara do que os outros métodos [ANTANI 02, CIOCCA 01, RUI 97]. Além disto, esta abordagem permite que seja definida uma composição de cores (conjunto simultâneo de cores), o que permite caracterizar adequadamente muitos objetos e cenas.

A abordagem utilizando RNAs requer que exista uma fase de treinamento da rede. Isto implica que a arquitetura do IMatch também possua um subsistema de geração da base de aprendizado de imagens e um subsistema de pesquisa (reconhecimento de imagens) na base. A Figura 1 apresenta o modelo esquemático do IMatch. No subsistema de geração da base de imagens prepara-se uma base que será utilizada para treinar a RNA. Esta base deve conter imagens de treinamento de duas categorias que denominaremos de: *Classe* e *Não-Classe*. As imagens da categoria *classe* são imagens-exemplo representantes do que se pretende que a RNA aprenda a reconhecer na procura (por exemplo, logotipos, latas de Coca-Cola, placas, sinalizações, etc). A categoria *Não-Classe* contém imagens-exemplo que não são relacionadas com a categoria *Classe* (tipicamente imagens-de-fundo) e servem para que a RNA aprenda a delimitar o espaço de procura para imagens da categoria *Classe*. Estas imagens de treinamento são extraídas de regiões retangulares de imagens que contenham a *Classe/Não-Classe*. Devemos definir o tamanho desta região de acordo com o tipo de imagens de treinamento disponível. Quanto maior for esta região em relação ao objeto a ser reconhecido, menor será a precisão de localização da *Classe* procurada porque irá ocorrer uma maior contribuição do fundo da imagem. Se a região escolhida for muito pequena em relação ao objeto perde-se em generalização dos atributos que compõem a *Classe*. No nosso caso escolhemos uma região de 64 x 64 *pixels*, para os experimentos que serão descritos a seguir.

A partir desta base de imagens de treinamento é então realizada uma extração de características principalmente referentes à cor utilizando o sistema de cor RGB ou HSV (Figura 1, Subsistema de Treinamento). Para cada imagem da base de treinamento aplica-se o método de redução de cor *Median Cut* [GOMES 94] que resulta numa paleta de N cores e, então se calcula o histograma de cores desta paleta. O resultado da extração de características é um vetor de tamanho igual a  $N * 4$ , onde N é a quantização de cores aplicada, sendo que para cada três conjuntos de componentes (RGB ou HSV) possuímos uma freqüência de ocorrência da cor (histograma). Este vetor é normalizado uniformemente para o intervalo [0, 1] e atribui-se um rótulo conforme a categoria da imagem de treino (“1 0” para *Classe* e “0 1” para *Não Classe*). Os vetores contendo as

componentes da paleta de cores, histograma e a classe correspondente são fornecidos à RNA. O treinamento da RNA é feito externamente ao IMatch utilizando o programa NeuSim [OSÓRIO 99] que simula o aprendizado de uma RNA MLP usando o algoritmo *Cascade-Correlation*. A generalização obtida pela RNA pode ser entendida como o reconhecimento da “assinatura” que caracteriza a coleção de imagens de treinamento. Esta assinatura aprendida a partir de uma coleção inicial de imagens é então reconhecida pela RNA usando as características extraídas da imagem-exemplo. Deste modo não estamos mais executando uma comparação 1-a-1, mas sim uma **comparação com o modelo generalizado de N imagens**, o que é um diferencial em relação às técnicas atuais para recuperação de imagens.



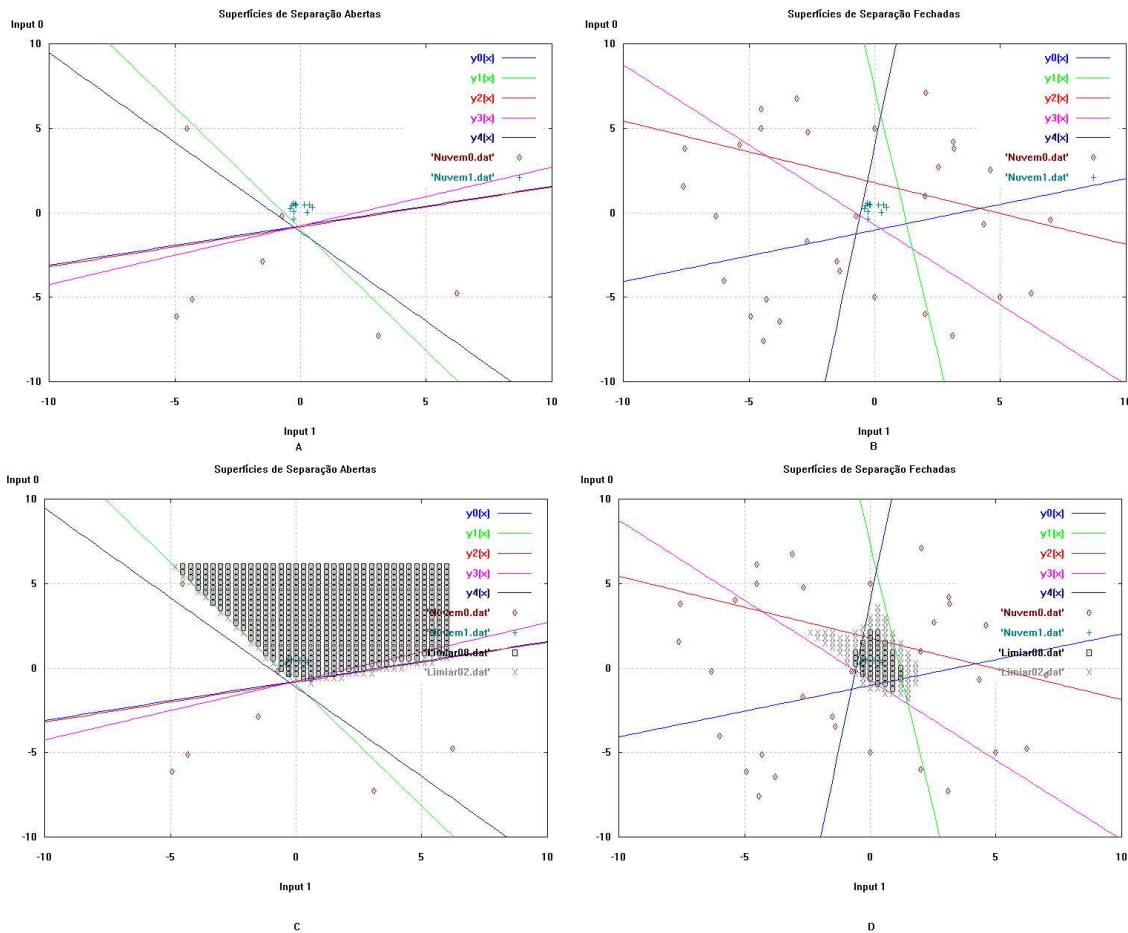
**Figura 1. Modelo esquemático dos subsistemas utilizados no IMatch.**

No subsistema de reconhecimento, o usuário informa ao programa uma imagem-exemplo e o programa irá percorrer esta imagem procurando a classe aprendida no treinamento. Neste reconhecimento, uma “janela deslizante”, que possui as mesmas dimensões da janela utilizada para o treinamento, varre a imagem-exemplo com uma sobreposição ajustável usando um deslocamento configurável pelo usuário (usualmente  $32 \times 32$  pixels). Para cada janela de varredura é aplicado o mesmo pré-processamento para extração de características que foi utilizado para o treinamento compondo o vetor de entrada que será passado para o SimNeu [OSÓRIO 99]. O SimNeu é um simulador de RNAs que usa uma rede previamente treinada e realiza a ativação da rede a fim de classificar os padrões que lhe são apresentados. A saída da RNA indicará se aquela região da imagem-exemplo é similar ao padrão que está sendo procurado (no caso, saída igual a 1 0). O programa permite ainda que se especifique um critério de limiar (*threshold criteria*) para ser aplicado a saída de ativação da rede.

O método de janela deslizante com sobreposição permite uma identificação mais precisa de qual região da imagem-exemplo possui o padrão procurado, o que é um diferencial em relação a outros sistemas CBIR que somente retornam uma coleção de imagens que atendem ao critério de pesquisa, sem destacar quais regiões da imagem foram responsáveis pela similaridade.

#### 4. Problema das Superfícies Fechadas de Classificação

Na tarefa de reconhecimento e recuperação de imagens é interessante que o método empregado recupere imagens com a menor diferença possível entre as características escolhidas da imagem-exemplo com as características das imagens da base de imagens procurando evitar ao máximo falsos reconhecimentos. No contexto da classificação de imagens utilizando RNAs isto significa evitar os falsos-positivos.



**Figura 2. Superfícies de Separação: A) Aberta; B) Fechada; C) Aberta com Limiar de Padrões; D) Fechada com Limiar de Padrões.**

Gori e Scarselli [GORI 98] apresentam um estudo sobre a utilização de RNAs do tipo MLP para a tarefa de classificação de padrões baseada na compreensão geométrica das superfícies de separação obtida pelo ajuste dos pesos dos neurônios da rede. Os autores apresentam uma análise de que, ao contrário dos bons resultados que vários trabalhos relatam, RNAs do tipo MLP utilizando critério de limiar, não são adequados

para serem utilizadas como classificadores quando os padrões apresentados à RNA não enquadram-se perfeitamente nas classes pré-definidas, portanto realizando uma rejeição confiável do padrão. Para ocorrer uma rejeição confiável de um padrão é necessário que as superfícies de classificação sejam **fechadas**. Ainda segundo Gori, em redes MLP com uma estrutura piramidal, nas quais a quantidade de neurônios da camada oculta é menor que quantidade de neurônios de entrada, **nunca** irão ocorrer superfícies fechadas de classificação. Caso a quantidade de neurônios da camada oculta seja maior que a quantidade de neurônios de entrada, não há garantia de ocorrer superfícies fechadas de classificação, e determinar esta condição é um problema da classe *NP-Hard*. Os autores do artigo sugerem que os casos de sucesso no uso de RNAs do tipo MLP para a tarefa de classificação se deve ao pré-processamento aplicado aos dados e a natureza destes dados.

Para exemplificar a proposta de Gori e Scarselli, treinou-se uma RNA do tipo MLP com a topologia de duas entradas, cinco neurônios na camada oculta e um neurônio de saída utilizando o algoritmo de aprendizado *back-propagation*. Foram utilizadas duas classes para os padrões de aprendizado que foram denominadas de *Nuvem0* e *Nuvem1*. Os padrões foram gerados aleatoriamente com uma dispersão maior para os padrões do rótulo *Nuvem0*. O objetivo é verificar se aumentando a quantidade de padrões *Nuvem0* **em torno** dos padrões *Nuvem1* consegue-se que a RNA determine superfícies de separação fechadas em torno dos padrões *Nuvem1*. Na Figura 2A tem-se um exemplo de uma superfície aberta de separação. As retas plotadas,  $y_0(x)$  a  $y_4(x)$ , representam as superfícies de separação determinadas pelos pesos dos neurônios da camada oculta e neste caso as suas intersecções não determinam um polígono em torno de todos os padrões *Nuvem1*. A Figura 2C apresenta sobreposto aos padrões de treinamento e às retas, os padrões de entrada para os quais a RNA treinada teve a sua saída ativada (padrões da classe *Nuvem1*), utilizando valor de limiar de 0.8 e 0.2. Podemos verificar neste exemplo que a RNA classifica os padrões de entrada pertencentes à região da classe *Nuvem0* como sendo pertencentes da classe *Nuvem1*, portanto sem executar uma rejeição confiável dos padrões da classe *Nuvem0*. As Figuras 2B e 2D demonstram como o aumento de padrões *Nuvem0* permitiu que a RNA ajustasse seus pesos de maneira mais adequada e determinasse uma superfície fechada de separação em torno dos padrões *Nuvem1*. As retas determinam um polígono bem próximo aos padrões *Nuvem1*. Na Figura 2D podemos observar que se utilizando um limiar de valor 0.8 a RNA irá executar uma rejeição confiável dos padrões *Nuvem0* e irá classificar corretamente padrões da classe *Nuvem1*.

A análise destes resultados nos levou a considerar que embora o trabalho de Gori indique que não há garantia de haver superfícies de classificação fechadas, elas são possíveis de se obter se conseguirmos fornecer padrões de treinamento que “cerquem” a região daqueles padrões que queremos que a RNA classifique como pertencentes à classe procurada na imagem. Fornecer estes padrões no contexto de recuperação de imagens é factível e isto nos levou a duas abordagens para tentar obter as superfícies de classificação fechadas que iremos abordar nas próximas seções.

#### **4.1. A influência das superfícies de classificação na recuperação de imagens**

O reconhecimento e recuperação imagens produz um vetor de entrada com uma variabilidade muito grande nos padrões que o compõe. O número de combinações de

cores e proporções em que elas ocorrem numa imagem é extremamente grande. Esta variabilidade é principalmente problemática nas regiões da imagem que não contém o padrão procurado (imagem de fundo). Isto implica que os padrões apresentados a RNA durante uma procura possuam grande probabilidade de se encontrarem fora do domínio dos padrões de treino. Como vimos na seção anterior, quando ocorre esta situação, RNAs do tipo MLP que não possuem regiões fechadas de classificação não poderão executar uma rejeição confiável, ocasionando assim um falso reconhecimento. Durante nossos experimentos com o IMatch foi verificada esta situação. Os resultados obtidos apresentaram mais falsos-positivos do padrão *Classe* do que o não reconhecimento do padrão *Classe*

## 4.2. Abordagem para Fechar as Superfícies de Classificação

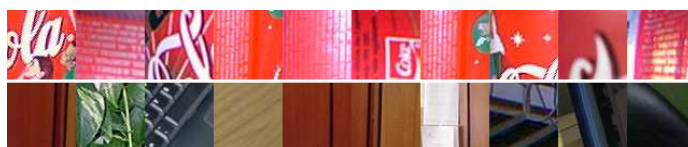
Nossa primeira abordagem (Regiões de Imagens) para se obter uma superfície de classificação fechada, ou quase fechada, foi utilizar regiões de algumas imagens classificadas erroneamente como pertencentes à categoria *Classe*, incluindo-as como exemplos de *Não-Classe* na base de treinamento e executando um novo treino da RNA. O sistema Ikona emprega uma abordagem similar de refinamento. Após a determinação dos métodos de extração e comparação de características, executam-se simulações de recuperação na base preparada. O usuário pode então executar um refinamento especificando nas imagens recuperadas, imagens como sendo exemplos **positivos** ou **negativos** que o sistema irá utilizar para melhor refinar as pesquisas futuras. O ponto negativo desta abordagem é obter uma RNA especializada para determinada base de imagem. Modificações no conteúdo da base de imagens tornariam necessária uma nova simulação de recuperação e talvez um novo refinamento.

A segunda abordagem (Nuvem de Não-Classe) para fechar as superfícies de classificação visou reduzir a especialização resultante de se utilizar regiões de imagens da base de pesquisa disponível e aumentar a automatização do processo. Para isto procuramos gerar automaticamente exemplos de *Não-Classe* que preenchessem o espaço de cores (conforme foi visto na Figura2D, a inclusão de exemplos de *Nuvem0 permitiu que a RNA ajustasse os pesos para determinar uma superfície fechada de classificação*) para serem utilizadas no treinamento da RNA junto com exemplos de *Não-Classe* provenientes de regiões de imagens. O método utilizado foi gerar novos padrões aleatoriamente utilizando uma distribuição uniforme para as  $n$  triplas de cores e  $n$  histogramas. Podemos citar como desvantagem o fato desta abordagem ser do tipo “força-bruta”.

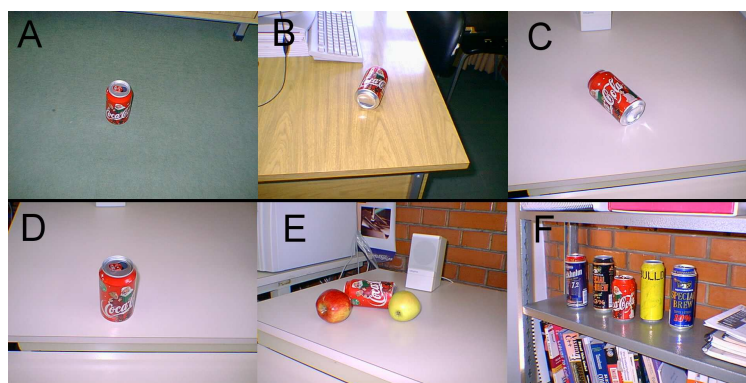
## 5. Resultados

Experimentos foram realizados utilizando uma RNA treinada para reconhecer latas de Coca-Cola. A base de treinamento foi confeccionada a partir de regiões de 64 x 64 pixels de imagens fotográficas obtidas com câmera fotográfica digital em condições não controladas de iluminação. Foram utilizadas 1000 regiões de imagens de latas de Coca-Cola de diferentes tipos, ângulos e aproximação como exemplos do padrão *Classe*, e 1000 regiões de imagens diversas (sem conter latas de Coca-Cola na sua composição) como exemplos do padrão *Não-Classe*. A Figura 3 apresenta algumas das imagens utilizadas como exemplo para os padrões. A Figura 4 mostra o conjunto de imagens-exemplo utilizadas na comparação das abordagens para o fechamento das superfícies de

classificação. As figuras 4A a 4D foram escolhidas para demonstrar a robustez do método quanto à rotação, variação de iluminação e escala do objeto procurado na cena. E as figuras 4E e 4F são cenas com uma grande variação de objetos e composição de cores. O sistema de cor utilizado foi o RGB com uma redução para 5 cores. O passo de varredura utilizado foi 32 x 32 pixels e um limiar de 0,96.



**Figura 3. Exemplo de imagens de treinamento dos padrões *Classe* (latas de Coca-Cola na primeira linha) e *Não-Classe* (segunda linha).**

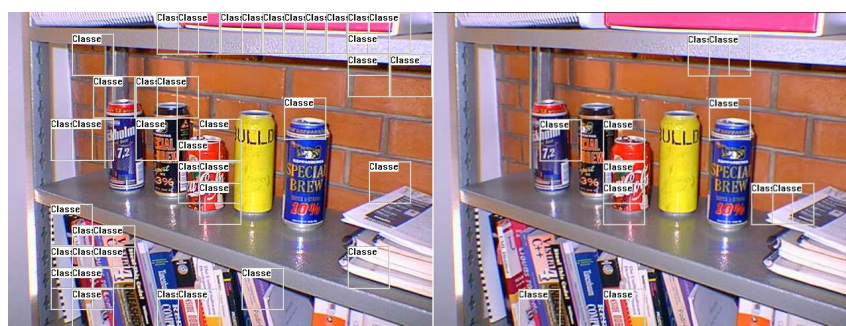


**Figura 4. Imagens utilizadas na comparação das abordagens para fechamento das superfícies de classificação.**

A Tabela 1 apresenta a contagem de regiões de 64 x 64 pixels que foram classificadas como sendo pertencente à categoria *Classe* utilizando como método para fechamento de superfícies de classificação a abordagem “Regiões de Imagens”. A tabela mostra a quantidade de erros (E) e acertos (A) para seis refinamentos (E1,A1 a E6,A6). Para as imagens 4A a 4D nos seis refinamentos o método não apresentou contribuições, mas também não causou falsos reconhecimentos.

A Figura 5 mostra um exemplo do reconhecimento para a imagem-exemplo F utilizando a abordagem “Regiões de Imagens”. A imagem da esquerda é o reconhecimento sem nenhum refinamento, executado apenas com o treinamento original da RNA, e a figura da direita após seis refinamentos, cada um deles incluindo algumas regiões que não foram classificadas corretamente como exemplos de *Não-Classe*. Observa-se que ainda existem regiões classificadas erroneamente, mas com uma redução significativa. A Tabela 2 apresenta a contagem de regiões de 64 x 64 pixels que foram classificadas como sendo pertencente à categoria *Classe* utilizando como método para fechamento de superfícies de classificação a abordagem “Nuvem de Não-Classe”. A cada refinamento foram adicionados com exemplos de *Não-Classe* na base de treinamento. A tabela mostra a quantidade de erros (E) e acertos (A) para seis refinamentos (E1,A1 a E6,A6). Verifica-se que este método é menos eficiente na redução dos falsos-positivos quando comparamos com a abordagem anterior.





**Figura 5. Exemplo de reconhecimento da imagem-exmplo F no primeiro e no último refinamento para a abordagem “Regiões de Imagens”.**

**Tabela 1. Contagem erros (E) e acertos (A) de regiões reconhecidas como *Classe* durante seis refinamentos utilizando a abordagem “Regiões de Imagens”.**

Imagem	E1	A1	E2	A2	E3	A3	E4	A4	E5	A5	E6	A6
A	0	1	0	4	0	0	0	2	0	0	0	3
B	0	2	0	4	2	3	0	2	0	1	0	1
C	0	11	0	10	0	6	0	7	2	4	0	9
D	0	7	0	10	0	8	0	10	0	7	0	11
E	10	7	11	7	14	7	9	5	7	7	1	4
F	35	5	38	5	39	3	17	5	17	3	8	3

**Tabela 2. Contagem erros (E) e acertos (A) de regiões reconhecidas como *Classe* durante seis refinamentos utilizando a abordagem “Nuvem de Não-Classe”.**

Imagem	E1	A1	E2	A2	E3	A3	E4	A4	E5	A5	E6	A6
A	0	1	0	1	0	0	0	1	0	1	0	0
B	0	3	0	2	0	5	0	3	0	2	0	2
C	0	6	0	8	0	9	1	9	1	7	0	10
D	0	10	1	9	0	9	1	11	1	8	0	8
E	14	7	8	6	11	6	9	8	18	6	7	6
F	35	3	25	5	39	5	32	3	31	5	28	3

## 6. Conclusões e Perspectivas

Com esta abordagem baseada em técnicas de Inteligência Artificial e aprendizado (*Machine Learning*) pretende-se melhorar a precisão da recuperação de imagens. O uso de um conjunto de imagens, no lugar de uma única imagem, nos permite “focar” melhor o processo de busca e recuperação de imagens em bases de dados. Apresentamos um embasamento teórico e uma verificação prática da necessidade de se considerar como uma RNA do tipo MLP irá executar uma classificação de padrões (e sua rejeição) baseado nas superfícies de separação. Também apresentamos duas abordagens para melhorar o ajuste das superfícies de separação no problema de recuperação e classificação de imagens baseados em atributos de cor através da inclusão de padrões de treinamento do tipo *Não-Classe*.

Baseado nos resultados obtidos verifica-se que para uma base de dados conhecida a melhor abordagem em termos de precisão é a “Regiões de Imagens” e que ainda é necessário melhorar a abordagem de “Nuvem de Não-Classe” no sentido de determinar uma melhor estimativa de quantos exemplos deverão ser gerados e como distribuí-los visando diminuir a quantidade de refinamentos e evitar que estes exemplos “invadam” a região dos padrões *Classe*.

## Referências

- ANTANI, Sameer; KASTURI, Rangachar; RAMESH, Jain. (2002) “A survey on the use of pattern recognition methods for abstraction, indexing and retrieval of images and video.”, *Pattern Recognition*, Amsterdã, Países Baixos, Elsevier Science B.V., n. 35, p. 945-965.
- BRUNELLI, Roberto; MICH, Ornella. (2000) “Image Retrieval by Examples.”, *IEEE TRANSACTIONS ON MULTIMEIA*, IEEE, v.20, n.3, p. 164-171.
- CIOCCA, Gianluigi; MARINI, Daniele; RIZZI, Alessandro; SCHETTINI, Raimondo; ZUFFI, Silvia. (2001) “On Pre-Filtering with Retinex in Color Image Retrieval”, *Internet Imagin II*, Proceedings of SPIE. V.4311, p. 140-147.
- GOMES, Jonas; VELHO, Luiz. *Computação Gráfica: Imagem*. (1994) IMPA/SBM, Rio de Janeiro, Brasil, p. 424.
- IKONA – IRNIA / IMEDIA Project Research Team. (2002) <http://www-rocq.inria.fr/imedia/ikona> (visitado em 10/10/2002).
- OSÓRIO, F. S.; AMY, B. (1999) “INSS: A hybrid system for constructive machine leaning.”, *Neurocomputing*. Amterdã, Elsevier Science B.V., v. 28, p. 191-205.
- FAHLMAN, Scott E.; LEBIERE, Christian (1991) “The Cascade-Correlation Learning Algorithm”, *Advances in Neural Information Processing Systems 3*, D. S. Touretzky (ed.), Morgan Kaufmann, Los Altos CA, p. 190-196.
- GORI, Marco; SCARSELLI, Franco. (1998) “Are Multilayer Perceptrons Adequate for Pattern Recognition and Verification ?”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v.20, n.11, Nov. 1998.
- Rui, Yong; HUANG, Thomas S.; CHANG, Shih-Fu. (1997) “Image Retrieval: Past, Present, and Future.”, *International Symposium on Multimedia Information Processing*, Taipei.
- SEBE, Nicu; LEW, Michael S. (2001) “Color-Based Retrieval.”, *Pattern Recognition Letters*, Amsterdã, Países Baixos, Elsevier Science B.V., n. 22, p. 223-230.
- SMITH, John R.; CHANG, Shih-Fu. (1996) “Tools and Techniques for Color Image Retrieval.”, *IS&T/SPIE In Symposium on Electronic Imaging: Science and Technology - Storage & Retrieval for Image and Video Databases IV*, San Jose, CA. v. 2670.