

Avaliação Estatística dos Descritores de Haralick na Detecção do Acidente Vascular Cerebral Isquêmico

Cecília Burle de Aguiar, Walisson da Silva Soares, Lucas Albuquerque de Lima e Carlos Danilo Miranda Regis

Resumo— Este artigo apresenta uma avaliação estatística dos descritores de Haralick utilizados para a detecção do acidente vascular cerebral isquêmico (AVCi). Com o objetivo de determinar quais os descritores essenciais para realizar essa detecção, podendo assim diminuir o número de descritores obtidos e utilizados no processamento. Dessa forma, diminuindo o tempo de processamento e o custo computacional. Utilizando o teste t foi possível identificar que os dois melhores descritores para essa aplicação são a soma das médias e a soma das variâncias.

Palavras-Chave— Descritores de Textura, Teste t , AVCi.

Abstract— This article presents one statistical evaluation on Haralick's textural features that were used to detection of ischemic stroke. Aiming determine which textural features are essentials to perform this detection, can thus reduce the number of features obtain and used in processing. In this way, reducing the processing time and the computational cost. Using t -test was possible to recognize that the two better Haralick's textural features were sum average and sum variance.

Keywords— Textural Features, t -Test, Ischemic Stroke.

I. INTRODUÇÃO

O acidente vascular cerebral (AVC) é, hoje, o maior causador de incapacidade funcional no ocidente, uma das causas mais comuns de disfunção neurológica na população adulta [1]. No Brasil e América do Sul, essa doença representa a principal causa de mortalidade e morbidade [2].

A classificação do AVC, quanto a sua causa, pode ser feita em dois tipos: hemorrágico (AVCh), decorrente da ruptura de vasos sanguíneos; e isquêmico (AVCi), decorrente da insuficiência de suprimento sanguíneo cerebral, podendo ser temporário (episódio isquêmico transitório, EIT) ou permanente [3].

Apesar dos sintomas apresentados nos dois tipos de AVC serem extremamente similares, o tratamento para cada um dos dois é completamente diferente. Assim, é necessária a visualização de imagens cerebrais para confirmação do diagnóstico de AVCi, que representa cerca de 80% de todos os casos de AVC [3].

O principal método diagnóstico de imagem nos casos de AVC é a tomografia computadorizada (TC). Como há pouca distinção das cores nas imagens da TC, em que mesmo um especialista com extensa experiência pode não identificar o AVCi, o uso da TC é para exclusão de outros diagnósticos e não para determinação do acometimento pelo AVCi [4].

Atualmente vem aumentando o número de sistemas automatizados de detecção dessas áreas [5]. Com o objetivo de

realizar a detecção automática das áreas de AVCi utilizando os descritores de Haralick, foi realizada nesse trabalho uma análise estatística de cada um desses descritores, a fim de utilizar àqueles que oferecem um maior percentual de confiança e precisão na detecção do AVC.

II. DESCRITORES DE HARALICK

Os descritores de Haralick podem ser utilizados para o cálculo da textura de uma imagem digital e são constituídos por 14 medidas estatísticas, ilustradas na Tabela I. Esta remete à distribuição espacial dos *pixels* que pertencem a uma textura pela construção da matriz de co-ocorrência.

A matriz de co-ocorrência é formada por elementos que relacionam a frequência que ocorre um determinado tom de cinza, X , em relação a outro, Y , que são separados por uma distância d_x de linhas e d_y de colunas. Pois a textura da imagem é intrínseca à relação de distribuição espacial média dos tons de cinza na imagem. E a descrição de textura baseia-se na repetição no aparecimento de um tom de cinza nessa textura.

A extração de características é realizada sobre os elementos da matriz de co-ocorrência fazendo-se diversos cálculos estatísticos. Essas características não só são utilizadas para diferenciação de uma região para outra, mas também são eficientes para análise e classificação de imagens, pois são uma assinatura do padrão definido pela região em estudo [6].

III. MATERIAIS E MÉTODOS

A. Base de Dados

A base de dados utilizada neste trabalho é composta por imagens de tomografia computadorizada de crânio de 20 pacientes. Essa base de dados foi disponibilizada pelo Núcleo de Estudos e Pesquisa em Imaginologia (NEPI) da Universidade Federal da Paraíba (UFPB). Todas as imagens seguem o padrão DICOM de imagens médicas e possuem dimensões de 512×512 . Cada exame de TC de cada paciente possui cerca de 20 imagens, com o espaçamento de 8 mm entre os cortes axiais da cabeça. Os laudos de cada paciente foram realizados por um médico especialista na área de radiologia, que são utilizados como padrão ouro de comparação.

Desses 20 pacientes, 10 deles não se apresentam com patologias, 6 com AVCi e 4 com outras patologias, como perda volumétrica, leucoaraiose, entre outras. Para o processamento, foi realizada a coleta dos pontos utilizados no cálculo dos descritores para o tratamento estatístico. Apesar de cada paciente possuir cerca de 20 cortes, foram selecionadas apenas 3 imagens de cada paciente para serem analisadas.

Cecília Burle de Aguiar, Walisson da Silva Soares, Lucas Albuquerque de Lima e Carlos Danilo Miranda Regis, Coordenação de Engenharia Elétrica, Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, João Pessoa-PB, Brasil, E-mails: burlececilia@gmail.com, danilo.regis@ifpb.edu.br.

B. Metodologia

O algoritmo desenvolvido neste trabalho é iniciado com a leitura da imagem no formato DICOM e a partir dela são realizadas as coletas de amostras em cada imagem. Em cada imagem são coletadas 5 amostras, nos exames com AVCi as amostras escolhidas são exatamente as da área acometida pela isquemia, enquanto nos outros pacientes são coletadas amostras do tecido cerebral normal.

Para cada uma das amostras coletadas na imagem são calculados todos os descritores de Haralick, após o cálculo dos 14 descritores para as 300 amostras presentes, esses valores dos descritores são organizados em 90 amostras de AVCi e 210, de tecido cerebral dentro da normalidade. Após a organização de todos os valores de descritores para cada uma das amostras aplica-se o teste t bicaudal, pois as amostras apresentam uma distribuição normal. O fluxograma da Figura 1 corresponde ao algoritmo desenvolvido neste trabalho.

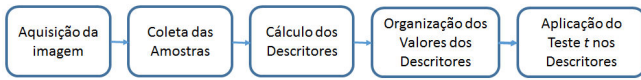


Fig. 1. Fluxograma do trabalho desenvolvido.

Para o tratamento estatístico foi utilizado o teste t para realizar a comparação bicaudal, utilizando as Equação 1 e 2. Nesse teste foi definido um intervalo de confiança de 95%, ou seja, o erro máximo atingido a ser aceito será de 5%. Neste trabalho, o teste t foi utilizado para identificar estatisticamente quais os descritores mais importantes na discriminação entre as amostras de tecido normal e com AVCi. Sendo útil lembrar que quanto mais próximo de zero o valor do teste t calculado, melhor o descritor correspondente será na distinção entre tecido sadio e com isquemia.

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\left(\frac{1}{n_1} + \frac{1}{n_2}\right) s_p^2}}, \quad (1)$$

em que

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}. \quad (2)$$

O valor da média de um descritor para amostras do AVCi e normais são representadas por \bar{x}_1 e \bar{x}_2 , respectivamente. Enquanto n_1 e n_2 representa o número de amostras em cada um dos grupos, e s_p^2 é a representação da variância ponderada. Os valores s_1^2 e s_2^2 representam o valor da variância no grupo do AVCi e no dos normais, respectivamente.

IV. RESULTADOS

Os resultados obtidos na aplicação do teste t aplicado para cada um dos descritores de Haralick são apresentados na Tabela I. É importante notar que quanto mais próximo de zero o valor do teste t do descritor correspondente, mais importante ele será na análise e menor a necessidade de retirá-lo do processamento.

É possível notar que apenas o valor correspondente ao descritor de correlação está acima dos 5% de erro, podendo assim ser desconsiderado como característica importante de distinção entre os *pixels* da área sadia e os da área isquemiada.

TABELA I

RESULTADO DO TESTE t DOS DESCRITORES.

| Descritores | Teste t |
|-------------------------|------------------------|
| Segundo Momento Angular | $1,19 \times 10^{-5}$ |
| Contraste | $9,87 \times 10^{-9}$ |
| Correlação | $5,41 \times 10^{-2}$ |
| Variância | $6,77 \times 10^{-9}$ |
| Homogeneidade | $5,69 \times 10^{-16}$ |
| Soma das Médias | $2,88 \times 10^{-86}$ |
| Soma das Variâncias | $3,26 \times 10^{-67}$ |
| Soma das Entropias | $3,97 \times 10^{-15}$ |
| Entropia | $1,19 \times 10^{-2}$ |
| Variância da Diferença | $2,21 \times 10^{-9}$ |
| Entropia da Diferença | $1,26 \times 10^{-19}$ |
| Energia | $7,72 \times 10^{-4}$ |
| Probabilidade Máxima | $3,71 \times 10^{-3}$ |
| Diferenças de Ordem K | $4,59 \times 10^{-4}$ |

Também observa-se que os dois valores significativamente menores da Tabela I são os dos descritores: Soma das Médias e Soma das Variâncias. Após essa observação foi elaborado um algoritmo de detecção de áreas de AVCi utilizando a combinação entre esses dois descritores apenas para teste e obteve-se uma acurácia de 76,19%. Assim, nesse processamento foram utilizados apenas esses dois descritores para se realizar a detecção das áreas de isquemia, minimizando o tempo de processamento sem perder desempenho.

V. CONCLUSÕES

Neste trabalho foi possível identificar estatisticamente que os dois descritores mais importantes na discriminação entre as amostras de tecido normal e com AVCi são a soma das médias e a soma das variâncias, quando analisados os descritores isoladamente. Podendo assim, diminuir o tempo de processamento apenas havendo o cálculo de dois descritores ao invés dos 14, obtendo-se uma boa acurácia.

É necessária a realização de uma análise estatística da combinação dos descritores, de 2 em 2, 3 em 3, entre outras, para assim, se realizar uma avaliação da detecção obtida utilizando a combinação de alguns descritores. A fim de analisar e melhorar as técnicas de detecção de AVCi utilizando os descritores de Haralick.

REFERÊNCIAS

- [1] C. André, *Manual de AVC*. Revinter, 1999.
- [2] P. M. Lavados, A. J. Hennis, J. G. Fernandes, M. T. Medina, B. Legetic, A. Hoppe, C. Sacks, L. Jadue, and R. Salinas, "Stroke epidemiology, prevention, and management strategies at a regional level: Latin america and the caribbean," *The Lancet Neurology*, vol. 6, no. 4, pp. 362–372, 2007.
- [3] G. A. Donnan, M. Fisher, M. Macleod, and S. M. Davis, "Stroke," *The Lancet*, vol. 371, no. 9624, pp. 1612–1623, Maio 2008.
- [4] E. Hudyma and G. Terlikowski, "Computer-aided detecting of early strokes and its evaluation on the base of ct images," in *Computer Science and Information Technology, 2008. IMCSIT 2008. International Multiconference on*. IEEE, 2008, pp. 251–254.
- [5] I. Rekić, S. Allasonnière, T. K. Carpenter, and J. M. Wardlaw, "Medical image analysis methods in mr/ct-imaged acute-subacute ischemic stroke lesion: Segmentation, prediction and insights into dynamic evolution simulation models. a critical appraisal," *NeuroImage: Clinical*, vol. 1, no. 1, pp. 164–178, 2012.
- [6] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *Systems, Man and Cybernetics, IEEE Transactions on*, no. 6, pp. 610–621, 1973.