

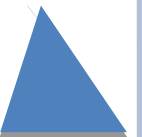


SEGMENTAÇÃO DE IMAGENS DE CÉLULAS CERVICAIS – BASE HERLEV

Processamento Digital de Imagens
Tarcísio Bruno C. Oliveira e
Jonathan Negreiros de Freitas

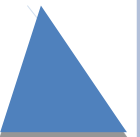
OBJETIVOS

- Avaliar quantitativamente o desempenho de cinco (05) diferentes métodos de segmentação. São eles:
 - Otsu;
 - Entropia de Renyi;
 - Crescimento de Região;
 - K-médias;
 - **Segmentação Fuzzy** (Fuzzy C-Means).
- Evidenciar a dificuldade ao segmentar imagens em problemas enfrentados na prática.



PROCEDIMENTO

- Não houve realização de pré-processamento nas imagens usadas;
- Os resultados obtidos nas métricas de validação serão, em alguns casos, valores baixos;
- As métricas de validação utilizados foram:
 - índice Dice;
 - Precision;
 - Recall;
 - False positive;
 - True positive;
- Avaliação das imagens a partir de composição das bandas por soma simétrica (Bloch, 1996) e média aritmética das bandas RGB.



MÉTODO: OTSU

- Algoritmo que assume que a imagem contém duas classes de *pixels*, seguindo uma distribuição bimodal (*pixels* de *background* e *foreground*), calculando um *threshold* que melhor separe as duas classes.
- Vantagem:
 - Possui boa performance se o histograma tem uma distribuição bimodal e, também, se possuir um **vale** bem definido entre dois **picos**;
- Desvantagem:
 - Caso a área do objeto em relação ao *background* seja consideravelmente menor, o algoritmo tende a piorar devido a ausência da bimodalidade.



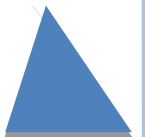
MÉTODO: ENTROPIA DE RENYI

- O algoritmo baseado na entropia de Renyi realiza a binarização de uma imagem através de duas distribuições probabilísticas (objeto e fundo do objeto)
- Vantagem:
 -
- Desvantagem:
 -



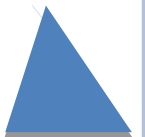
MÉTODO: CRESCIMENTO DE REGIÃO

- Algoritmo que define regiões pré-selecionadas através de uma semente (*seed*) inicial. A abordagem ocorre verificando se o *pixel* vizinho à semente pertence, ou não, a sua região, e, em caso positivo, tal pixel é incorporado a essa região.
- Vantagem:
 - Regiões com mesmas propriedades são corretamente separadas;
 - Imagens com bordas bem definidas resultam em melhores segmentações;
- Desvantagem:
 - Sensível a ruído;
 - Aplicado de maneira local, desconsiderando a visão geral do problema;
 - Computacionalmente custoso.



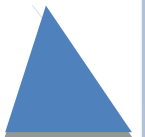
MÉTODO: K-MÉDIAS

- Algoritmo que realiza o particionamento (separação) em “k” *clusters* para um conjunto de dados. A afiliação (atribuição) de um dado pertencente a um *cluster*, é definido pela distância deste dado para um centroide K_n qualquer. A escolha de um centróide para cada *cluster* é definido por uma *seed*, e esta *seed* varia a cada nova execução do algoritmo.
- Vantagem:
 - Fácil implementação;
 - Rápida execução;
 - Versões variadas são criadas para atingir propósitos diferentes.
- Desvantagem:
 - A escolha aleatória de uma *seed* pode levar a resultados diferentes a partir de uma mesma imagem. Isto se acentua quando o conjunto de dados possui ruído.



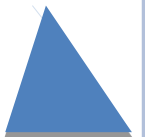
MÉTODO: Fuzzy C-Means

- Algoritmo semelhante ao K-Médias. O que diferencia este algoritmo, é o fato de ele utilizar um parâmetro chamado “*membership*” que indica o quanto um *pixel* pode estar relacionado a um particionamento, em detrimento de outro. No K-Médias, um pixel só pode pertencer exatamente à um *cluster*.
- Vantagem:
 - Fácil implementação;
 - Entrega bons resultados para dados correlacionados;
- Desvantagem:
 - Para melhores resultados, é necessário um maior número de iterações;
 - Especificação do número de *clusters* à priori.



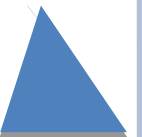
MÉTRICAS DE VALIDAÇÃO

- Este termo é utilizado para verificar a qualidade de segmentação de uma imagem;
- Por exemplo: em um cenário onde é feita a comparação de uma imagem segmentada com o “ground truth”, os resultados obtidos revelarão o quanto a segmentação foi efetiva em realizar a separação daquilo que se quer verificar em relação ao “ground truth”.



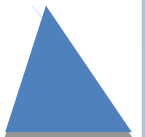
MÉTRICAS DE VALIDAÇÃO: DICE

- Conhecido como “coeficiente de similaridade”;
- Método estatístico usado para comparar a similaridade entre duas amostras;
- Verifica quão similar um conjunto de dados é em relação à outro.



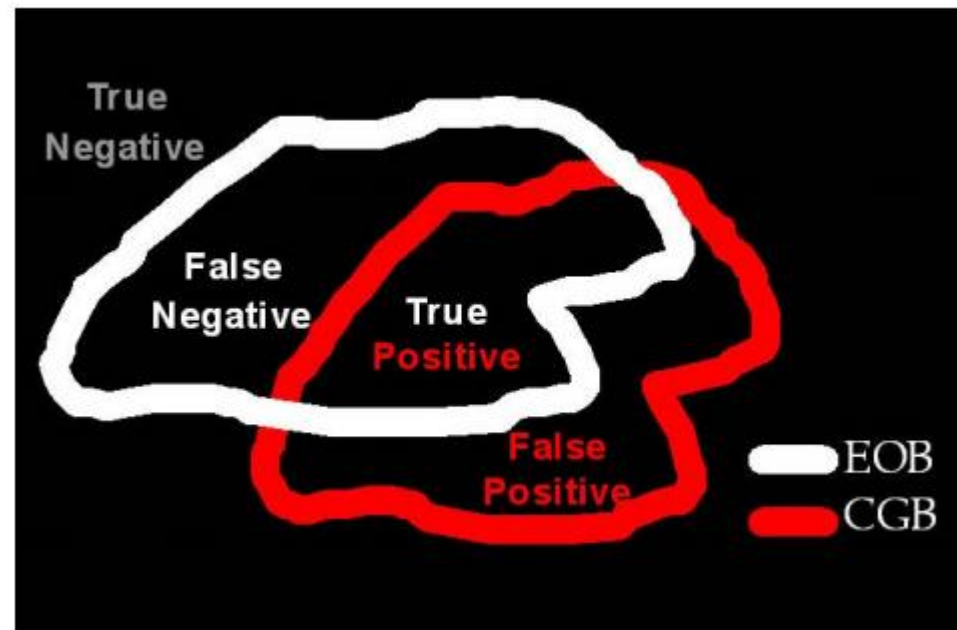
MÉTRICAS DE VALIDAÇÃO: PRECISION e RECALL

- O uso das métricas de PRECISION e RECALL são baseadas nos limites de região da imagem para avaliar a consistência de uma segmentação;
- **Recall:** é definido como a proporção de pixels de contorno do “*ground truth*” que foram detectados com sucesso pela segmentação;
- **Precision:** é definido como a proporção de pixels de contorno da segmentação que correspondem aos pixels de contorno do “*ground truth*”.

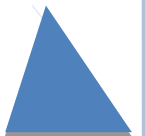


MÉTRICAS DE VALIDAÇÃO: FALSE POSITIVE e FALSE NEGATIVE

- **FALSE POSITIVE:** *pixels* segmentados **erroneamente** como *foreground* (o objeto da imagem em estudo);
- **FALSE NEGATIVE:** *pixels* segmentados **erroneamente** como *background* (fundo da imagem em estudo);



- EOB (*expert-outlined boundaries*) como “*ground truth*”;
- CGB (*computer-generated boundaries*).



RESULTADOS: Carcinoma

- Média aritmética das bandas RGB

| Algoritmo | Dice | Erro |
|----------------|--------|--------|
| | | |
| Otsu | 0.8070 | 0.0290 |
| K-Means | 0.4944 | 0.0917 |
| Renyi | 0.7501 | 0.0369 |
| Region growing | 0.4908 | 0.0478 |
| Fuzzy C-Means | 0.8133 | 0.0346 |



RESULTADOS: Carcinoma

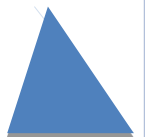
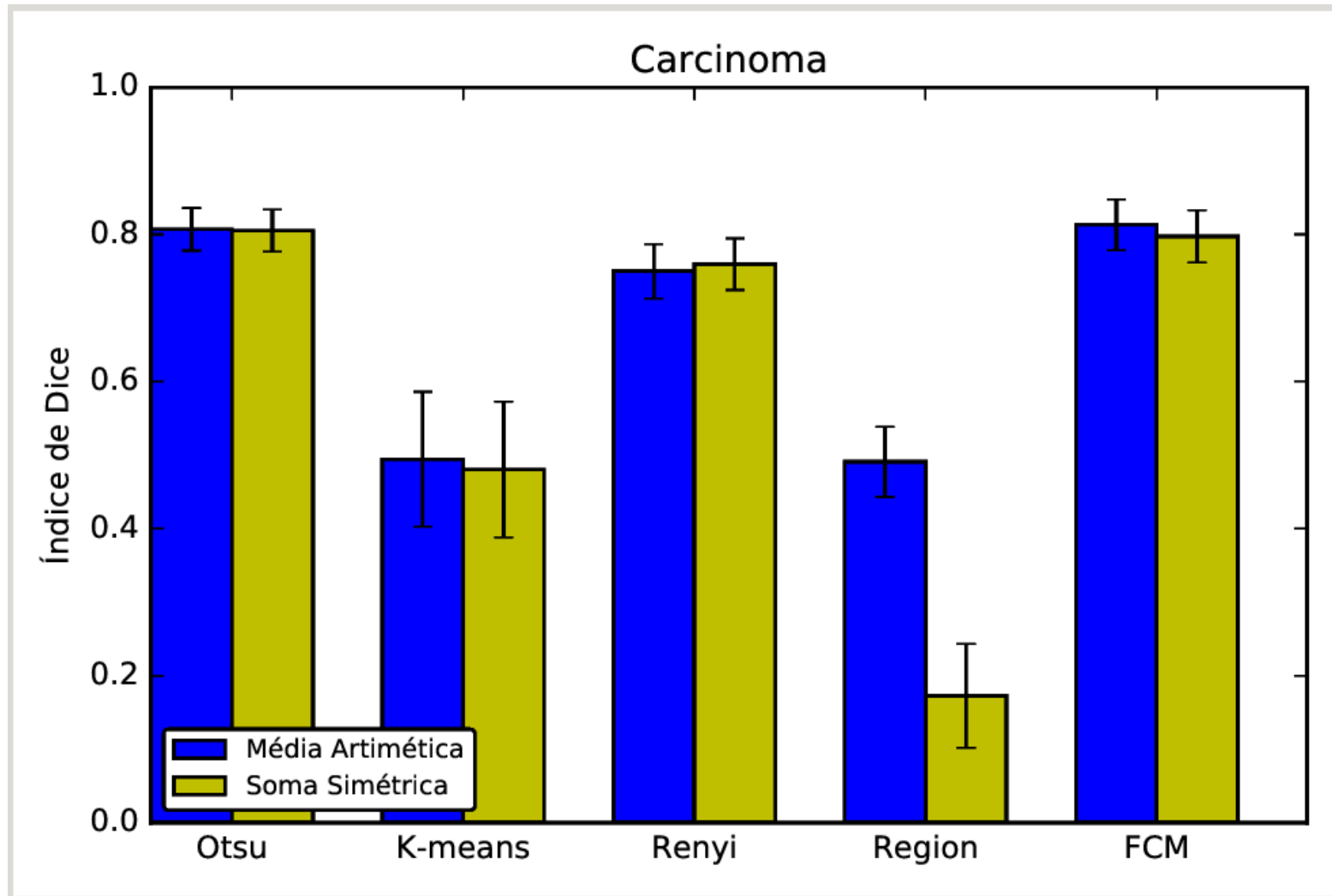
- Imagem por composição das bandas por soma simétrica (Bloch, 1996).

| Algoritmo | Dice | Erro |
|----------------|--------|--------|
| | | |
| Otsu | 0.8053 | 0.0285 |
| K-Means | 0.4804 | 0.0924 |
| Renyi | 0.7595 | 0.0350 |
| Region growing | 0.1729 | 0.0705 |
| Fuzzy C-Means | 0.7972 | 0.0355 |



RESULTADOS: Carcinoma

- Gráfico



RESULTADOS: Moderate Dysplastic

- Média aritmética das bandas RGB

| Algoritmo | Dice | Erro |
|----------------|--------|--------|
| | | |
| Otsu | 0.6183 | 0.0400 |
| K-Means | 0.2491 | 0.0657 |
| Renyi | 0.5990 | 0.0432 |
| Region growing | 0.2820 | 0.0395 |
| Fuzzy C-Means | 0.7683 | 0.0300 |



RESULTADOS: Moderate Dysplastic

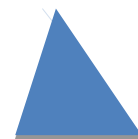
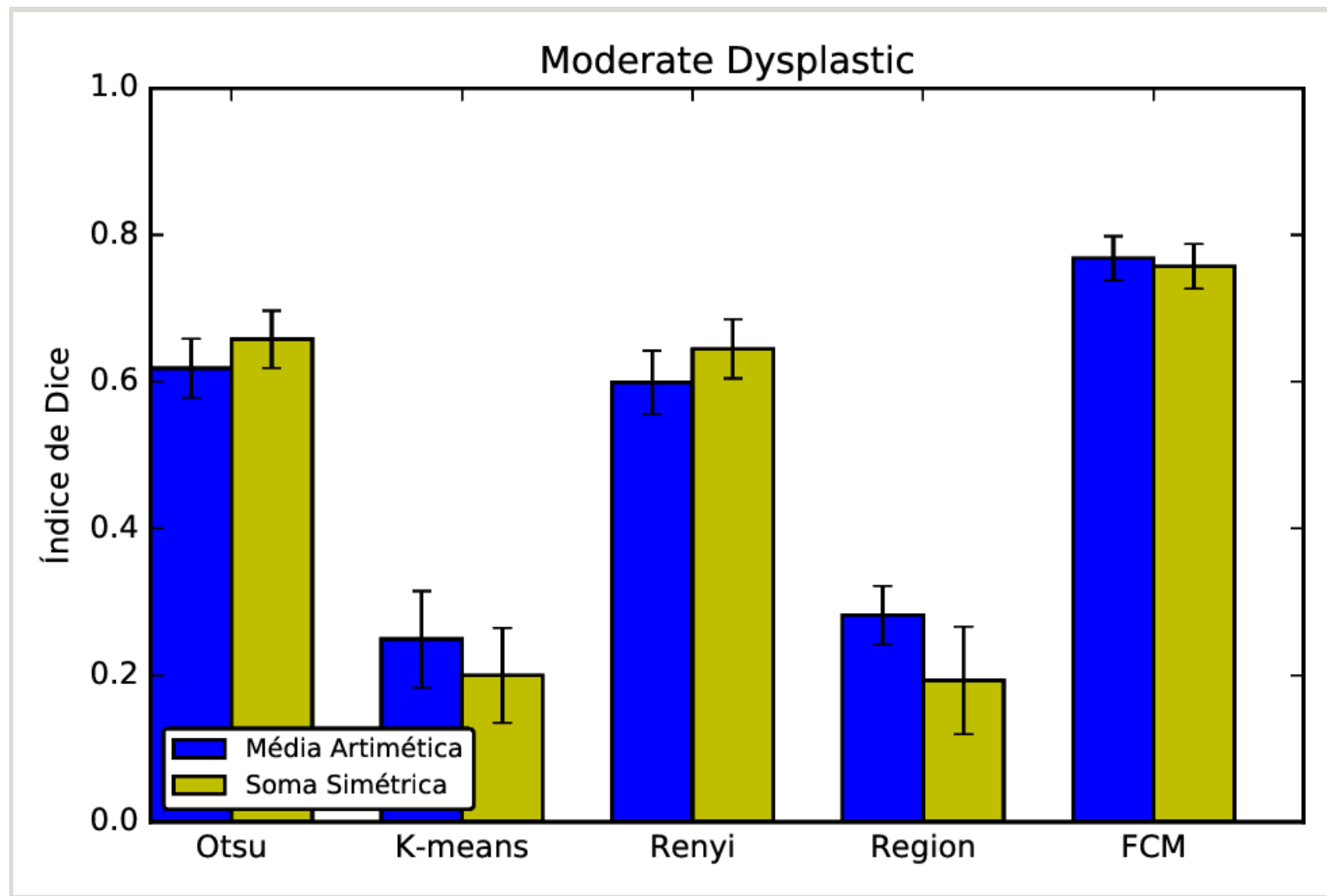
- Imagem por composição das bandas por soma simétrica (Bloch, 1996).

| Algoritmo | Dice | Erro |
|----------------|--------|--------|
| | | |
| Otsu | 0.6579 | 0.0388 |
| K-Means | 0.2001 | 0.0644 |
| Renyi | 0.6447 | 0.0402 |
| Region growing | 0.1931 | 0.0734 |
| Fuzzy C-Means | 0.7572 | 0.0306 |



RESULTADOS: Moderate Dysplastic

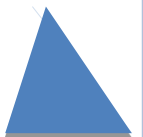
- Gráfico



RESULTADOS: Normal Intermediate

- Média aritmética das bandas RGB

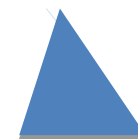
| Algoritmo | Dice | Erro |
|----------------|--------|--------|
| | | |
| Otsu | 0.1236 | 0.0371 |
| K-Means | 0.0360 | 0.0078 |
| Renyi | 0.4973 | 0.0482 |
| Region growing | 0.0049 | 0.0033 |
| Fuzzy C-Means | 0.2942 | 0.0491 |



RESULTADOS: Normal Intermediate

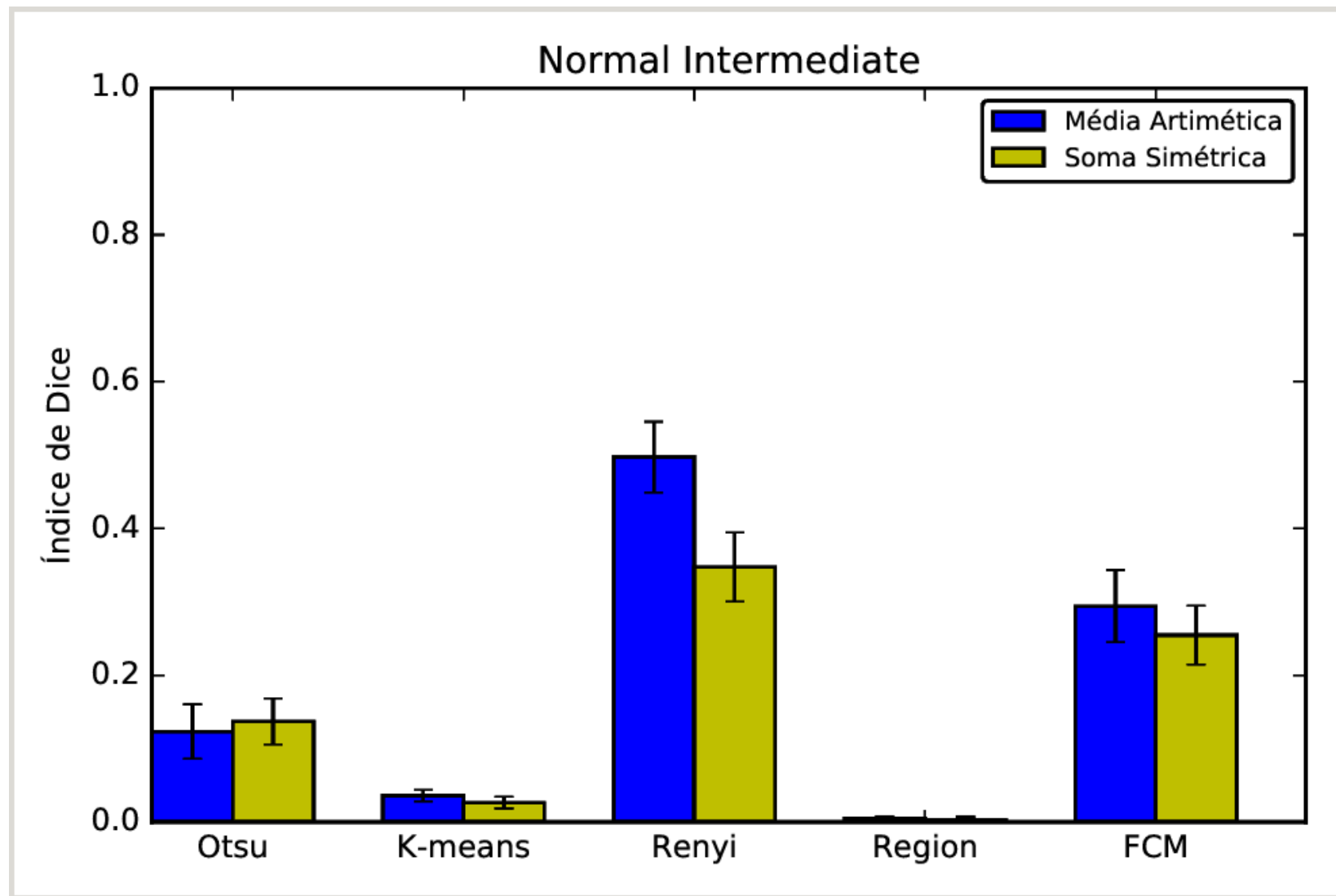
- Imagem por composição das bandas por soma simétrica (Bloch, 1996).

| Algoritmo | Dice | Erro |
|----------------|--------|--------|
| | | |
| Otsu | 0.1372 | 0.0316 |
| K-Means | 0.0266 | 0.0079 |
| Renyi | 0.3475 | 0.0473 |
| Region growing | 0.0032 | 0.0043 |
| Fuzzy C-Means | 0.2550 | 0.0402 |



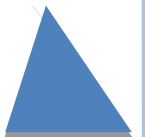
RESULTADOS: Normal Intermediate

- Gráfico



REFERÊNCIAS

- GONZALEZ, Rafael C.; WOODS, Richard E. **Processamento de Imagens Digitais**. Edgard Blucher, 2000.
- LU, Zhi et al. Evaluation of Three Algorithms for the Segmentation of Overlapping Cervical Cells. 2016.
- ZIJDENBOS, Alex P. et al. Morphometric analysis of white matter lesions in MR images: method and validation. **IEEE transactions on medical imaging**, v. 13, n. 4, p. 716-724, 1994.
- BYRD, Kenneth A.; ZENG, Jianchao; CHOUIKHA, Mohamed. A Validation model for segmentation algorithms of digital mammography images. **Journal of Applied Science & Engineering Technology**, v. 1, 2007.



REFERÊNCIAS

- ESTRADA, Francisco J. **Advances in computational image segmentation and perceptual grouping**. Chapter 5. 2005. Tese de Doutorado. University of Toronto.
- PESHKOV, Klim; PRÉVOT, Laurent. Segmentation evaluation metrics, a comparison grounded on prosodic and discourse units. In: **LREC**. 2014. p. 321-325.

