CATALOGAÇÃO DE FLORES ATRAVÉS DA ANÁLISE DE HISTOGRAMAS

Luan Pereira [1]

Martina Rodrigues [2]

Sidinei Ghissoni [3]

Resumo:

O reconhecimento de imagens vem ganhando grande espaço em diversas áreas e mercados diferentes, pois as técnicas de reconhecimento automático são utilizadas para tarefas gerais. O trabalho apresenta uma proposta de implementação de um algoritmo para a catalogação de especies de flores. O processo ocorre com a analise dos histogramas das imagens, que vão mostrar se as amostras são coerentes com o banco de dados, a partir é feita uma analise estatística nos resultados, que são apresentados ao usuário. A escolha de flores para a simulação do algoritmo, deu-se por conta da cloração distinta de cada espécie de flor, tendo em mente é claro que a analise de histograma necessita de uma amostra que não esteja poluída com outras cores, por isso foram escolhidas flores de espécies totalmente diferentes. Portanto, o trabalho visa demonstrar os aspectos, negativos e positivos, do método empregado para a catalogação das imagens. Visamos montar uma algoritmo genérico, que sera usado em situações e usos diversos.

Palavras-chave: analise, flores, histograma, algoritmo

Modalidade de Participação: Iniciação Científica

CATALOGAÇÃO DE FLORES ATRAVÉS DA ANÁLISE DE HISTOGRAMAS

[1] Aluno de graduação. luancenation122@gmail.com. Autor principal

[2] Aluno de Graduação. mtinarodrigues99@gmail.com. Co-autor

[3] Docente. sidinei.ghissoni@gmail.com. Orientador

CATALOGAÇÃO DE FLORES ATRAVÉS DA ANÁLISE DE HISTOGRAMAS

1 INTRODUÇÃO

O reconhecimento de imagens vem ganhando grande espaço em diversas áreas e mercados diferentes, pois as técnicas de reconhecimento automático são utilizadas para tarefas gerais (BARLA, 2003). Neste trabalho, propomos o uso específico de um algoritmo de comparação de histogramas para a catalogação de certos tipos de flores mais comuns. Foram utilizadas três principais técnicas para a catalogação das imagens: Intersecção de histogramas, análise de variância dos resultados e o método estatístico de Tukey Kramer (DRISCOLL, 1996) para uma precisão maior nos resultados.

Além de estudar as técnicas de catalogação de imagem, conseguimos também facilitar o trabalho de catalogação de certas espécies de plantas, tendo em vista que foi obtido uma precisão de cerca de 71,4% em testes primários com uma baixa população amostral. Sendo assim, pode-se considerar a usabilidade do algoritmo mesmo com baixo número amostral.

2 METODOLOGIA

Para ser feita a catalogação das imagens, antes de mais nada, faz se necessário uma amostra do usuário, para que possa ser feita a comparação com o banco de imagens do algoritmo. Essa imagem não deverá conter uma cor de fundo muito saturada em contraste com a cor da própria imagem, pois a análise de histograma depende basicamente das cores do objeto a ser analisado(ADRIAN, 2014). A Figura 1 apresenta o modelo de fluxograma do funcionamento do algoritmo.

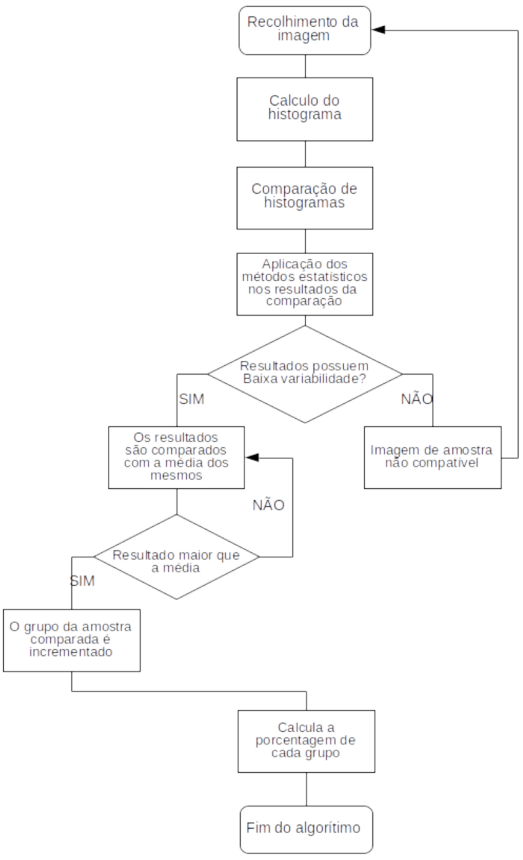

Figura 1: Fluxograma do funcionamento do algoritmo.

Quando uma imagem analisada é muito discrepante das demais, ela é automaticamente descartada antes mesmo de passar pelos métodos estatísticos, o que implica em um aumento

considerável no desempenho do algoritmo. Porém, pelo baixo número de amostras usadas nos testes, essa medida pode ser considerada negativa, pois a diminuição das amostras causa um decrescimento na precisão dos testes.

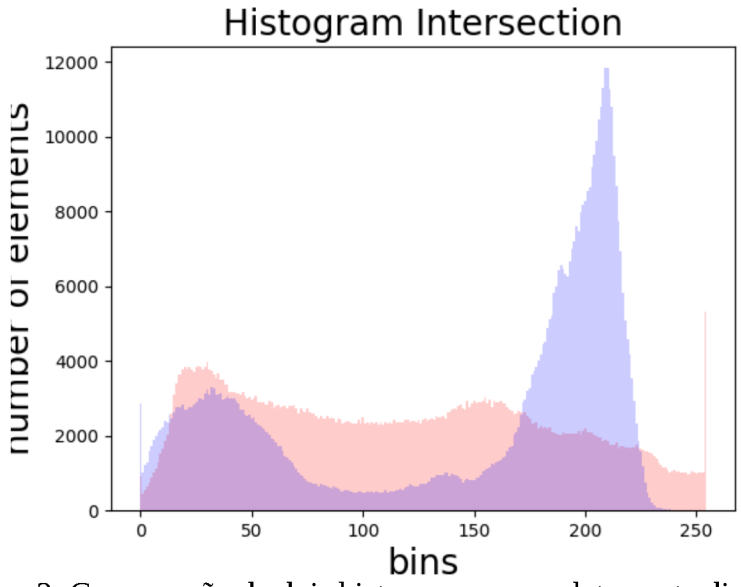

Figura 2: Comparação de dois histogramas completamente discrepantes.

Como pode ser visualizado na Figura 2, a comparação obteve apenas 58% de intersecção (área mais escura) entre os histogramas, o que implica em uma amostra discrepante do comparativo com o banco de dados.

É importante ressaltar também que a resolução das imagens de controle e da imagem amostral deve ser a mesma, pois o histograma é uma relação entre número de pixeis e a saturação dos mesmos.

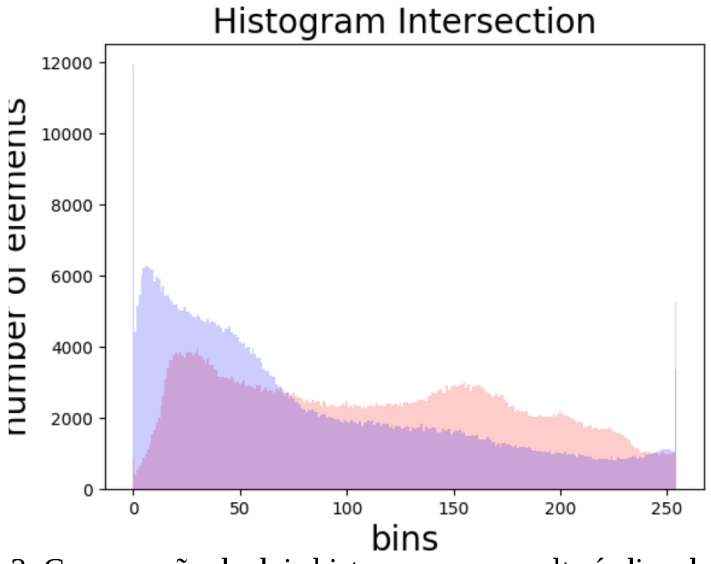

Figura 3: Comparação de dois histogramas com alto índice de intersecção.

Já na figura 3, conseguimos observar uma iteração de 77% entre as imagens analisadas, o que indica que a amostra corresponde ao grupo da imagem comparada.

Logo que a imagem passa por essa comparação de histogramas, é feita a aplicação dos métodos estatísticos no conjunto de resultados das comparações, os quais só vão ser considerados relevantes se possuírem uma baixa taxa de variabilidade.

2.1MÉTODO ANOVA

O método de teste de variância foi utilizado no algoritmo para medir a variabilidade das diferenças entre a amostra e o banco de dados. Com a aplicação do método, conseguimos estipular que uma amostra só vai ser relevante se a variância dos valores obtidos dentro de um grupo for de baixa variabilidade.

Os testes de variância das amostras não conseguem mostrar exatamente qual dos grupos é o mais coerente com a amostra do usuário, porém, ele é necessário para definir se um conjunto de dados é ou não suficiente para chegar a um resultado satisfatório. Com a utilização desse método podemos aumentar a porcentagem de acerto do algoritmo(DRISCOLL, 1996).

3 RESULTADOS E DISCUSSÃO

Mesmo com a aplicação dos métodos estatísticos, não foi possível obter uma eficácia totalmente satisfatória, pois obteve-se um resultado médio de 71,4% com uma amostragem de 80 exemplares por grupo, que é uma amostragem considerada baixa e 14 simulações. No entanto, pelo baixo custo de processamento, cerca de 15 segundos para execução total do algoritmo com 560 iterações entre imagens, e devido à velocidade conseguida nas comparações, sendo em média 0,02 segundos por comparação, podemos considerar uma alta usabilidade para o software mesmo em sistemas mais simples e compactos.

	Grupos de flores %							
Amostras	Classe 1	Classe 2	Classe 3	Classe 4	Classe 5	Classe 6	Classe 7	Resultado
Classe 1	**21,21**	17,5	18,5	5	13,4	7,7	16,5	Correto
Classe 1	18,02	11,5	20	7,5	**20,4**	9,9	12,5	Errado
Classe 2	**19,6**	17,14	14,6	5,7	18,21	10,35	14,28	Errado
Classe 2	22,5	**23**	16,7	3,27	8,36	6,18	20	Correto
Classe 3	10,5	3,9	22,1	16,3	**24,2**	15,9	6,8	Errado
Classe 3	17,9	10,6	**22**	7,3	21	8,3	12,9	Correto
Classe 4	10,9	4,4	16,4	**21,23**	20,8	15,7	10,2	Correto
Classe 4	13,7	3,6	17,32	**19,13**	17,68	18,4	10,1	Correto
Classe 5	18,9	12,54	13,89	11,1	**19,66**	8,1	15,6	Correto
Classe 5	13,6	4,4	19,11	16,38	**22,52**	17,06	8,82	Correto
Classe 6	19,33	11,33	**20**	8,66	15,66	11,66	13,33	Errado
Classe 6	15,18	5,9	19,4	14,5	11,22	**21,78**	11,88	Correto
Classe 7	19,7	22,49	16,26	3,11	10,38	5,19	**22,9**	Correto
Classe 7	21,84	17,06	12,96	4,43	10,58	8,53	**24,57**	Correto

Tabela1: Resultado das simulações.

Como pode ser visto na Tabela 1, as imagens do usuário estão contidas na coluna “Amostras”, enquanto o “Grupos de flores” mostra a porcentagem que o algoritmo extraiu de cada amostra para uma classe.

Além do baixo custo computacional, também há um mínimo custo de hardware, pois o usuário só necessita de uma câmera convencional para a retirada da amostra. Todo processamento e tratamento da imagem é feita pelo próprio software, o que melhora sua usabilidade e acesso.

4 CONSIDERAÇÕES FINAIS

Até o presente momento, não foi elaborada interface gráfica para o software, que conta somente com comandos via linha de código. Como perspectiva futura será feita a implementação de uma interface mais acessível ao usuário, e também o uso de redes neurais artificiais para a análise mais precisa dos dados recolhidos.

Pretende-se também utilizar o algoritmo para uso mais amplos de reconhecimento e catalogação de imagens, tendo em vista as abrangentes utilidades do mesmo como uso para reconhecimento de objetos e padrões de cores, catalogação de doenças cutâneas, entre enumeras outras aplicações. O software será implementado em dispositivos *mobile*, pois possui um baixo grau de processamento computacional, o que implica em uma abrangência maior de usos para o mesmo.

REFERÊNCIAS

A. Barla, F. Odone and A. Verri, "Histogram intersection kernel for image classification,"*Proceedings 2003 International Conference on Image Processing (Cat. No.03CH37429)*, Barcelona, Spain, 2003, pp. III-513.
doi: 10.1109/ICIP.2003.1247294
URL:http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1247294&isnumber=27938

HISTOGRAM Calculation. Disponível em: <https://docs.opencv.org/2.4/doc/tutorials/imgproc/histograms/histogram_calculation/histogram_calculation.html>. Accesso em: 08 maio 2018.

ROSEBROCK, Adrian.**Clever Girl**: A Guide to Utilizing Color Histograms for Computer Vision and Image Search Engines. 2014. Disponível em: <https://www.pyimagesearch.com/2014/01/22/clever-girl-a-guide-to-utilizing-color-histograms-for-computer-vision-and-image-search-engines/>. Acesso em: 08 maio 2018.

ROSEBROCK, Adrian.**How-To**: 3 Ways to Compare Histograms using OpenCV and Python. 2014. Disponível em: <https://www.pyimagesearch.com/2014/07/14/3-ways-compare-histograms-using-opencv-python/>. Acesso em: 08 maio 2018.

WADE C. Driscoll, Robustness of the ANOVA and Tukey-Kramer statistical tests, Computers & Industrial Engineering,
Volume 31, Issues 1–2, 1996, Paginas 265-268.