

Segmentação de Imagens de Folhas

Relatório Técnico

Squad 1 - Atlântico Bootcamp - 2022

Membros do Squad 1:

Emerson Veras
Felipe Araújo
Paulo Vinicius
Rômulo Castro

1. Apresentação do Problema

O presente relatório tem por objetivo descrever a construção de um sistema de tratamento computacional de imagens capaz de segmentar fotos de folhas de árvores diversas. As etapas de construção do sistema envolvem a aquisição de imagens, definição de um padrão ouro de segmentação, avaliação de algoritmos selecionados pela equipe e seleção do melhor com base na métrica *Intersection Over Union* (IOU).

Os links do código fonte do projeto no Github e a pasta no google drive com os arquivos gerados pelo projeto são encontrados nos links abaixo:

Github: <https://bit.ly/3yiHU3C>

Google drive: <https://bit.ly/3xWWM6f>

2. Aquisição de Imagens

Para a realização das imagens iniciais, um padrão para aquisição das fotos foi definido na tentativa de obter melhores resultados. O padrão para aquisição das imagens utilizadas como entradas do processo seguiu as especificações abaixo:

Fundo: branco (recomendado folha de A4)

Distância: uma caneta bic de distância - 15 cm

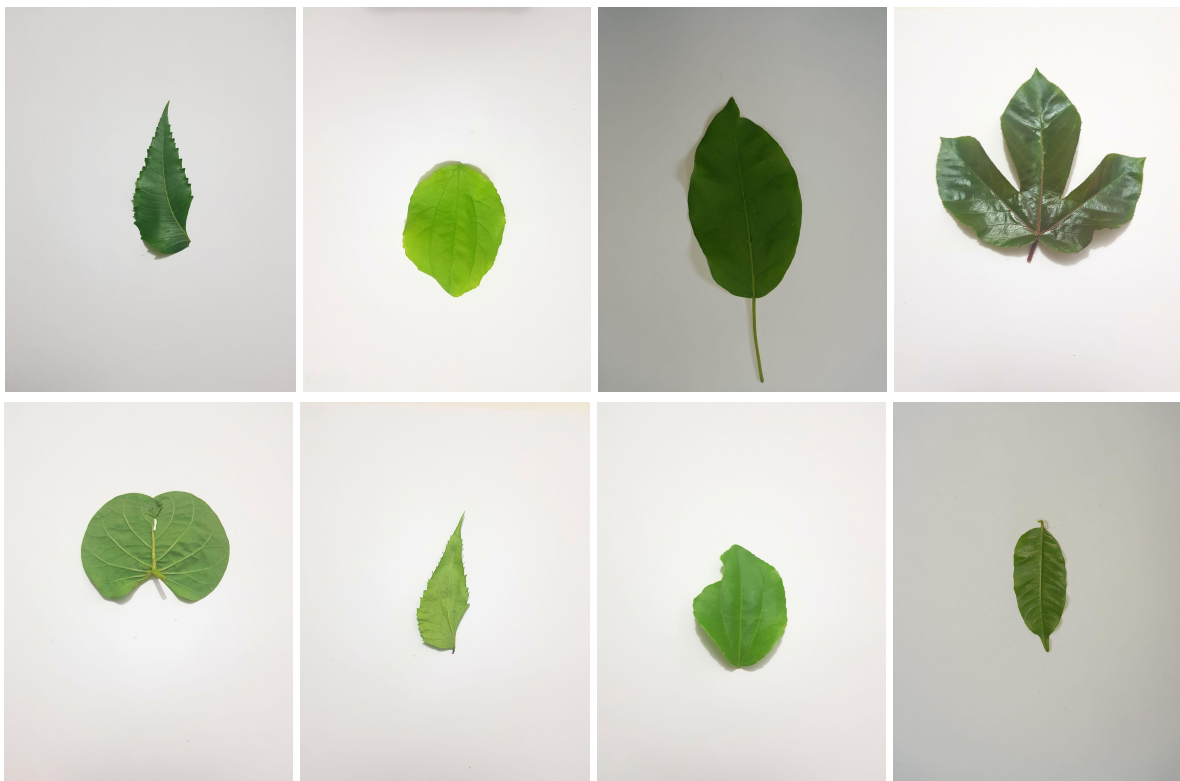
Formato: 4:3

Iluminação: artificial (fluorescente) sem sombra

Resolução da câmera: 4MP - 2880 x 2180

Número de imagens: 100 imagens

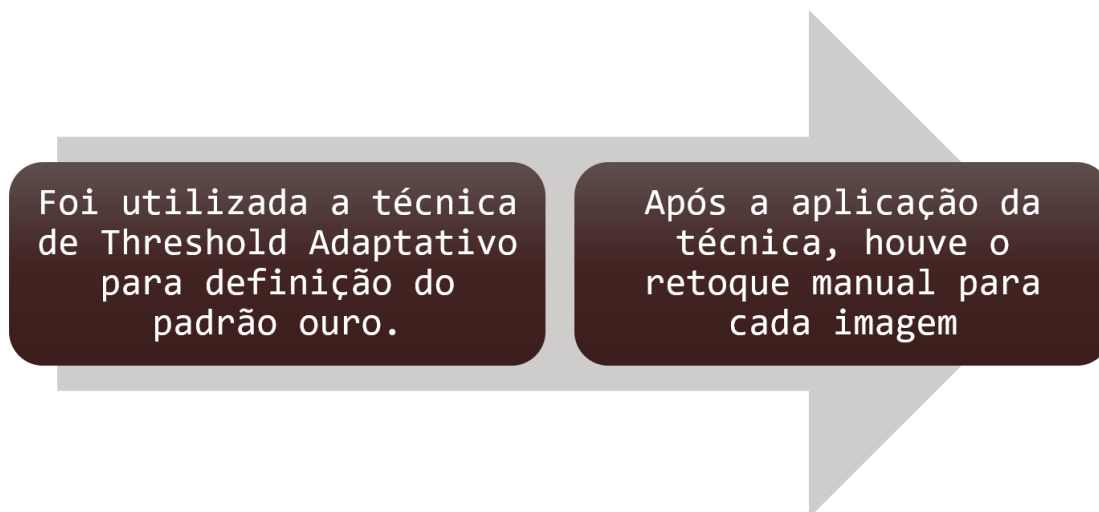
Figura 1 - exemplos das imagens geradas.



3. Definição do Padrão Ouro

O padrão ouro é a melhor segmentação possível para uma determinada imagem, e serve como referência para avaliação dos algoritmos propostos. Para o presente projeto, o padrão ouro foi definido através do pipeline apresentado na Figura 2.

Figura 2 - pipeline para definição do padrão ouro.



A primeira etapa foi a aplicação da técnica de Threshold adaptativo, inserida no conjunto de bibliotecas do Scikit-Image para linguagem Python. Na sequência, retoques manuais foram realizados nas imagens de padrão ouro que ainda apresentavam, dentro da região segmentada, pontos ou regiões em preto. Foi utilizada a ferramenta Paint do Windows para a tarefa. Na pasta do Google drive do projeto, obtêm-se todas as imagens padrão ouro realizadas. Abaixo na Figura 3 são apresentados alguns exemplos:

Figura 3 - exemplos de imagens padrão ouro



4. Aplicação de três (03) técnicas de segmentação

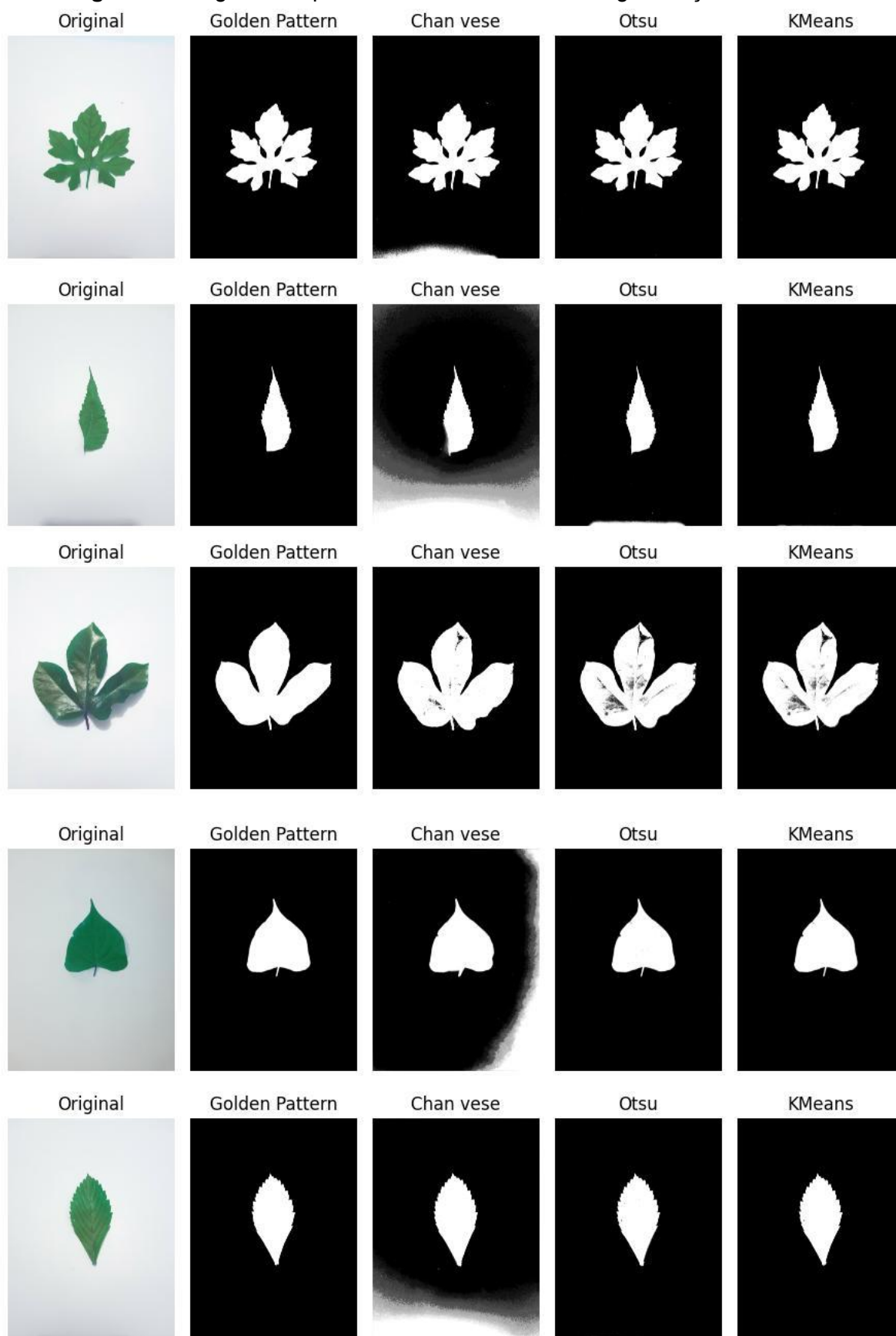
As técnicas utilizadas no projeto para segmentação foram as seguintes:

- Threshold de Otsu
- Chan-Vese
- Kmeans

4.1 Comparação entre as técnicas

Abaixo, algumas imagens comparativas das três técnicas:

Figura 4 - imagens comparativas das 3 técnicas de segmentação utilizadas.



A partir de uma observação visual e análise dos valores obtidos na aplicação da métrica IOU, percebe-se que a técnica k-means apresenta melhores resultados. No entanto, como mostram muito bem os outros squads, outras técnicas ainda podem ser aplicadas para obtenção de melhores resultados.

5. Extração de atributos e criação do CSV

Os atributos solicitados para cada imagem padrão ouro foram os seguintes:

- Referência do arquivo
- Dimensões iniciais
- Maior largura
- Maior comprimento
- Área

O arquivo gerado está na raiz do nosso projeto no GITHUB com o nome *golden_pattern_dims.csv*. De forma geral, os dados coletados no arquivo estão abaixo, na tabela 1.

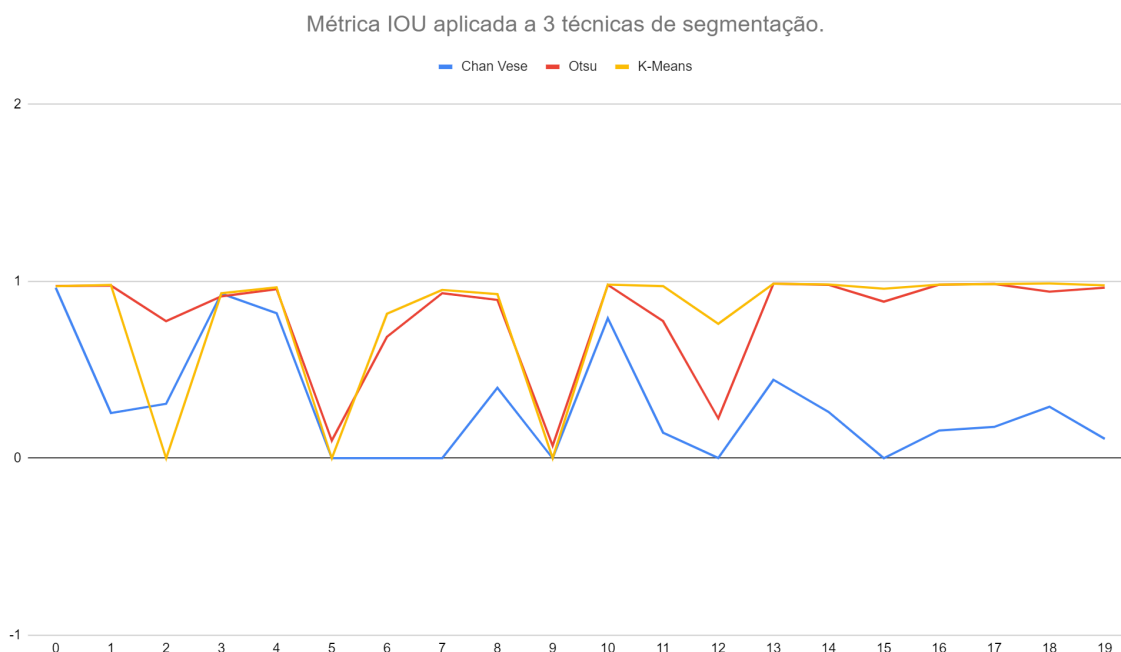
Tabela 1 - dados armazenados em csv com os atributos das imagens padrão ouro. Valores em pixels.

File_name	Orig_height	Origl_width	Max_height	Max_width	Area
IMG_20220604_110757.jpg	2448	3264	1886	2277	2173492
IMG_20220604_103838.jpg	3264	2448	996	510	386769
lmg (40).jpg	4032	3024	2136	682	904139
lmg (1).jpg	4032	3024	1963	1906	2135490
lmg (17).jpg	4032	3024	2131	1493	1269541
IMG_20220604_124653.jpg	2448	3264	412	950	168399
IMG_20220604_082935.jpg	2448	3264	645	1022	454627
IMG_20220604_082817.jpg	2448	3264	1492	440	495725
lmg (23).jpg	2880	2160	1307	497	373937
IMG_20220604_082908.jpg	2448	3264	334	616	150825
lmg (7).jpg	4032	3024	1581	1466	1334062
lmg (15).jpg	4032	3024	1560	524	458779
IMG_20220604_125220.jpg	3264	2448	1765	752	596317
IMG_20220604_103728.jpg	3264	2448	1147	929	602669
lmg (30).jpg	2880	2160	1187	450	313669
IMG_20220604_082844.jpg	2448	3264	760	879	472714
IMG_20220604_103748.jpg	3264	2448	1465	325	331303
IMG_20220604_103851.jpg	3264	2448	1059	507	402728
lmg (41).jpg	4032	3024	1663	791	767046
lmg (20).jpg	4032	3024	1390	503	425857

6. Aplicação da métrica IOU e geração de CSV

A técnica de avaliação *INTERSECT OVER UNION* (IOU) ou INTERSECÇÃO SOBRE A UNIÃO demonstra a medida de acurácia entre um objeto detectado e um conjunto específico. Abaixo na figura 5, uma representação dos valores obtidos somente com a aplicação das técnicas citadas na seção 4.

Figura 5 - gráfico dos valores das métricas IOU.



Observa-se pela figura 5 a disparidade obtida nos gráficos. Como razão para tal resultado elenca-se a não uniformidade na aquisição das imagens e a falta de uso de outras técnicas para aperfeiçoamento do resultado. Abaixo na tabela 2, estão os dados estatísticos calculados da tabela 1:

Tabela 2 - dados estatísticos obtidos da aplicação da métrica IOU.

	Chan Vese	Otsu	K-Means
Máx. value	0.96271740730000	0.9860119307	0.9873055299
Mín. value	0.00005517389259	0.0701133873	0
Average	0.30242270750000	0.7985055731	0.8052757960
Std. deviation	0.32565714400000	0.3000970965	0.3519402724

Na tabela 2, percebe-se no desvio padrão que a disparidade se apresenta. Aproximadamente 35% de desvio padrão entre os valores do dataset é algo que merece atenção e a criação de um pipeline com técnicas que possam melhorar/reduzir esses valores.

7. Seleção da Melhor Técnica

Para seleção da melhor técnica, foi utilizada a média dos valores *IOU* obtidos para as 20 imagens do conjunto de teste, conforme tabela 2. De acordo com este critério, a melhor técnica foi o K-Means.

8. Discussão dos Resultados

As três técnicas utilizadas apresentaram resultado satisfatório para a tarefa proposta. Paradoxalmente, apesar de apresentar a melhor média para a métrica *IOU*, a segmentação usando K-Means também resultou em algumas imagens com segmentação completamente errada, com apenas 1 grande cluster aparente.

Percebeu-se também que a segmentação pelo método de Chan-Vese, por ter sido desenvolvida para tratar especificamente de objetos com contornos não muito definidos, apresentou maior sensibilidade às variações de iluminação, por vezes tratando sombras como objetos de interesse.

Observando os gráficos das métricas *IOU* e as imagens de exemplo na figura 4, nota-se que os três métodos seguem uma tendência geral: quando um deles performa mal para determinada imagem, os demais igualmente apresentam desempenho ruim, e vice-versa, ou seja, quando um apresenta um bom resultado para certa imagem, os outros igualmente conseguem bom desempenho na segmentação da mesma.

9. Conclusões e Possíveis Melhorias

As três técnicas escolhidas para segmentação apresentaram resultados semelhantes na maioria dos casos, e possivelmente, para um sistema a ser utilizado em uma aplicação do mundo real. Os três poderiam ser empregados em conjunto, combinando resultados da mesma forma que ocorre nos problemas de classificação em que se utiliza a técnica de comitê ou *ensemble learning*.

Notadamente, a iluminação foi o fator que mais afetou os resultados finais, e, a depender da aplicação, este é um fator sobre o qual se pode ter maior ou menor controle, podendo inclusive serem especificados padrões mais estritos de forma a controlar sua variabilidade.

Por fim, mesmo na técnica que apresentou melhor métrica, nota-se alguns pixels classificados erroneamente. Na pasta *raw_segmented_best* encontram-se as imagens segmentadas com a melhor técnica e nas mesmas dimensões da imagem original, e, aplicando-se zoom, é possível identificar tais pixels. Uma possibilidade de melhoria para o resultado final seria aplicar alguma técnica morfológica, tal com crescimento de regiões ou erosão, a fim de eliminar esses elementos classificados incorretamente.