AVALIAÇÃO SISTEMÁTICA DE TÉCNICAS CLÁSSICAS E BASEADAS EM APRENDIZADO DE MÁQUINA PARA A IDENTIFICAÇÃO E CONTAGEM DE OBJETOS DISJUNTOS EM IMAGENS

Caetano Vendrame Mantovani (PIC/UEM), Marcelo de Gomensoro Malheiros (Orientador). E-mail: ra135846@uem.br, mgmalheiros@uem.br

Universidade Estadual de Maringá, Departamento de Informática, Maringá, PR.

Área e subárea do conhecimento: Ciência da Computação / Metodologia e Técnicas da Computação.

Palavras-chave: Visão Computacional; Processamento de Imagens; Framework de Avaliação.

RESUMO

A identificação e a contagem de objetos de uma mesma categoria em imagens é um problema clássico de Visão Computacional, resolvido tanto por técnicas clássicas explícitas quanto por abordagens de Aprendizado de Máquina, cada uma com vantagens e custos distintos. Este projeto de Iniciação Científica tem como objetivo construir um *framework* computacional para avaliação sistemática dessas técnicas aplicadas à contagem de objetos disjuntos, isto é, sem sobreposição, comparando diferentes abordagens sob os mesmos critérios de acurácia, tempo de execução e uso de memória. Na etapa atual, foram implementadas e avaliadas cinco técnicas clássicas de segmentação: limiarização adaptativa, detecção de bordas, transformada de Hough para círculos, limiarização por histerese e segmentação por *watershed.* Estas técnicas foram aplicadas sobre um *dataset* de fotografias de moedas brasileiras sem sobreposição, usando um mesmo protocolo automatizado de avaliação. Os resultados mostram diferenças relevantes entre as técnicas: a segmentação por *watershed* obteve a maior acurácia (92%) e o menor erro médio, enquanto a limiarização por histerese foi a mais rápida e mais leve em memória, porém a menos precisa (71%); a transformada de Hough, apesar de acurada (90%), consumiu bem mais memória que as demais. Esses resultados evidenciam, de forma concreta, os *trade-offs* entre acurácia, custo computacional e uso de memória que motivam o projeto, e validam o *framework* como ferramenta útil para comparar objetivamente diferentes técnicas. Os próximos passos incluem incorporar técnicas de Aprendizado de Máquina ao mesmo *framework* e estender a avaliação a outros conjuntos de dados, reais e sintéticos.

INTRODUÇÃO

Identificar e contar objetos de uma mesma categoria em uma imagem é uma tarefa clássica de Visão Computacional, com aplicações que vão da análise de imagens médicas à automação industrial, sem que exista uma técnica de uso universal (GONZALEZ; WOODS, 2018). Técnicas clássicas – limiarização, detecção de bordas, morfologia matemática – têm processo explícito e baixo custo computacional, o que as torna atrativas para ambientes controlados ou de borda (BAYGIN et al., 2018; MOON et al., 2022), já o Aprendizado de Máquina tende a ser mais robusto e genérico, ao custo de maior complexidade de treinamento e inferência (RANJAN et al., 2021).

Este projeto de iniciação científica tem como objetivo geral construir um framework computacional para comparar sistematicamente técnicas clássicas e de Aprendizado de Máquina aplicadas à contagem de objetos, avaliando objetivamente os *trade-offs* entre acurácia, custo computacional e uso de memória. A etapa atual concentra-se em objetos disjuntos, isto é, sem sobreposição entre instâncias, reduzindo o problema a uma versão mais simples e tratável, na mesma linha da abordagem explícita adotada por Peiter et al. (2025).

MATERIAIS E MÉTODOS

A avaliação foi conduzida sobre um conjunto de 6.021 imagens anotadas do Brazilian Coin Detection Dataset (MONEDA; YONEKURA; GUEDES, 2020), contendo fotografias de moedas brasileiras (5, 10, 25 e 50 centavos além de 1 real) sem sobreposição e sobre fundo uniforme, com anotações *labelme* (círculos e polígonos) que fornecem a contagem real por imagem, entre 1 e 14 moedas.

O framework foi implementado em Python (scikit-image, SciPy e NumPy), organizando cada técnica em um *notebook* que compartilha um módulo de avaliação comum. Esse módulo mede, de forma padronizada, quatro critérios – acurácia, erro médio absoluto, tempo médio de execução por amostra e pico médio de memória – implementando o objetivo de comparar as técnicas sob os mesmos critérios (GONZALEZ; WOODS, 2018).

Foram implementadas cinco técnicas clássicas de segmentação, todas seguidas por uma limpeza morfológica (dilatação, preenchimento de buracos e erosão) e contagem por componentes conexos: (i) limiarização adaptativa de Sauvola; (ii) detecção de bordas de Canny; (iii) transformada de Hough para círculos, sobre bordas pré-processadas; (iv) limiarização por histerese sobre o gradiente de Sobel; e (v) segmentação por *watershed* a partir de marcadores de intensidade. Os parâmetros de cada técnica foram ajustados manualmente por inspeção visual em uma pequena amostra antes da avaliação quantitativa sobre todo o conjunto.

RESULTADOS E DISCUSSÃO

A comparação entre as cinco técnicas foi consolidada usando os mesmos critérios de avaliação, o que permite comparação direta entre elas. A Tabela 1 resume esses resultados sobre as imagens do conjunto anotado.

Tabela 1 – Comparação das cinco técnicas clássicas avaliadas sobre as 6.021 imagens do subconjunto anotado.

Técnica	Acurácia	Erro médio	Tempo por amostra (s)	Memória (MB)
Limiarização (Sauvola)	80,11%	0,3298	0,1	17,2
Bordas (Canny)	88,82%	0,1758	1,0325	14,4
Hough (círculos)	90,45%	0,1192	0,6430	68,4
Histerese + morfologia	71,23%	0,6384	0,0622	12,0
Watershed (regiões)	92,69%	0,0959	0,4166	12,3

A segmentação por *watershed* obteve a maior acurácia e o menor erro médio absoluto, seguida da transformada de Hough; a histerese foi a técnica mais rápida e mais econômica em memória, porém a menos acurada, enquanto o Hough, apesar de acurado, consumiu de quatro a seis vezes mais memória que as demais – reflexo de avaliar o acumulador para toda a faixa de raios considerada. Com esses testes, é possível perceber que essas diferenças evidenciam os *trade-offs* entre acurácia, tempo e memória que motivam a proposta do framework (PEITER et al., 2025). A Figura 1 ilustra esse pipeline em uma imagem-exemplo com seis moedas, todas corretamente identificadas: imagem original (a), máscara segmentada (b) e contagem final por região conectada (c).

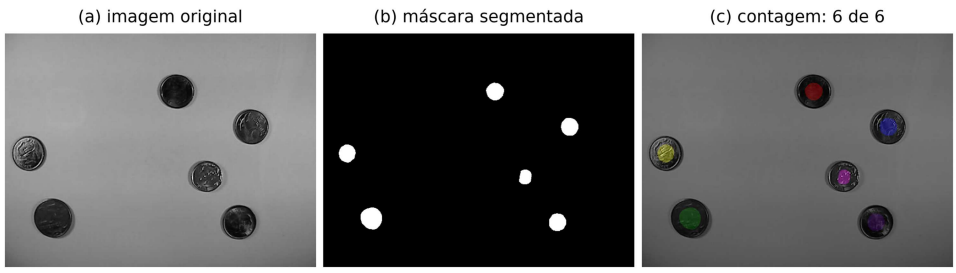


Figura 1 – *Pipeline* de contagem aplicado a uma imagem de exemplo: (a) foto original; (b) máscara segmentada após limpeza morfológica; (c) contagem final por região conectada (técnica de *watershed*).

CONCLUSÕES

O *framework* proposto já cumpre seu papel de comparar objetivamente diferentes técnicas sob os mesmos critérios, mesmo restrito a técnicas clássicas e a um dataset de objetos disjuntos, com diferenças relevantes de acurácia, tempo e

memória entre elas – por exemplo, a histerese é preferível com poucos recursos computacionais, e o *watershed* quando a acurácia for prioritária. Os próximos passos são incorporar Aprendizado de Máquina ao mesmo *framework* (RANJAN et al., 2021) e estender a comparação a outros *datasets*.

REFERÊNCIAS

BAYGIN, M. *et al.* **An Image Processing based Object Counting Approach for Machine Vision Application**. 2018. Disponível em: https://arxiv.org/abs/1802.05911. Acesso em: 19 ago. 2026.

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. 4. ed. [S.l.]: Pearson, 2018.

MONEDA, L.; YONEKURA, D.; GUEDES, E. **Brazilian Coin Detection Dataset**. [S.l.]: IEEE Dataport, 2020. DOI: 10.21227/kj4c-y809. Disponível em: https://ieee-dataport.org/open-access/brazilian-coin-detection-dataset. Acesso em: 19 ago. 2026.

MOON, J. *et al.* Smart count system based on object detection using deep learning. **Remote Sensing**, v. 14, n. 15, 2022. ISSN 2072-4292. Disponível em: https://www.mdpi.com/2072-4292/14/15/3761. Acesso em: 19 ago. 2026.

PEITER, A. F. *et al.* Circle-based object counting using image analysis. *In*: **2025 38th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)**. [S.l.: s.n.], 2025. p. 1–6.

RANJAN, V. *et al.* Learning to count everything. *In*: **2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2021. p. 3393–3402.