AVALIAÇÃO DA SEGMENTAÇÃO DIGITAL DE IMAGENS APLICADA À CONTAGEM DE SEMENTES DE SOJA (*Glycine max*)

Maria Iza Silva de Castro[1], Maria Eduarda Lopes Fernandes[2], Caio Nepomuceno Verdeiro[2], Flaviane Mendonça Ambrozim[2], Poliany de Oliveira Barbosa[2], Patricia Alvarez Cabanez[3], Lucimara Ribeiro Venial[2],

[1]Universidade Federal do Ceará/Departamento de Fitotecnia, Avenida Mister Hull, s/n, Pici - 60455-760 - Fortaleza-CE, Brasil, mariaizasc2011@alu.ufc.br.
[2]Universidade Federal do Espírito Santo/ Programa de Pós-Graduação em Energia (PPGEN)/Laboratório LaPAQui/ Centro Universitário Norte do Espírito Santo (CEUNES), Rodovia BR-101 Norte, Km 60, Litorâneo - 29932-540 - São Mateus-ES, Brasil, mariaeduardalopesfernandes582@gmail.com, caioverdeiro@gmail.com, flaviane.ambrozim@ufes.br, poliany.barbosa@ufes.br, lucimara.venial@ufes.br.
[3]Universidade Federal Rural do Rio de Janeiro, Avenida Lourival Martins Beda, 1200, Parque Varanda do Visconde - 28178-000- Campo dos Goytacazes-RJ, Brasil, patriciacabanez@gmail.com.

Resumo
Este trabalho teve como objetivo avaliar a segmentação digital aplicada à contagem de sementes de soja (*Glycine max*) e interpretar os descritores morfométricos e colorimétricos extraídos de cinco lotes, utilizando um script desenvolvido em MATLAB sem toolboxes proprietárias. Foram digitalizadas por escâner, sobre fundo azul padronizado, 20 imagens (5 lotes × 4 repetições de 100 sementes). A contagem automática coincidiu com a contagem manual de referência em todas as imagens (2.000 sementes, erro absoluto nulo). A área média das sementes variou de 4.141 a 4.380 px entre lotes, amplitude da mesma ordem da variação entre repetições de um mesmo lote (CV de 1,4% a 10,6%). O canal B apresentou a maior amplitude relativa entre lotes (82,4 a 98,6), com CV entre repetições de até 3,0%. A segmentação digital apresentou concordância integral com a contagem manual na condição padronizada avaliada, constituindo evidência preliminar de sua aplicabilidade como ferramenta auxiliar para a contagem e a descrição morfométrica e colorimétrica de lotes de sementes.
Palavras-chave: Visão computacional. Sementes de soja. Contagem automática. Fenotipagem digital. Processamento de imagens.

Área do Conhecimento: Ciências Agrárias/Agronomia.

Introdução

A soja é uma das principais culturas agrícolas do país, e a qualidade do lote de sementes utilizado no plantio impacta diretamente a produtividade final (Scheeren *et al*., 2010). De acordo com Zu *et al*., (2025), a contagem de sementes é uma etapa recorrente em análises de pureza, germinação e vigor, tradicionalmente conduzida manualmente por analistas treinados, processo lento e sujeito à variabilidade entre avaliadores.

Ferramentas de visão computacional têm sido propostas para reduzir o tempo e a subjetividade associados à contagem manual, com desempenho competitivo frente a ela em diversas espécies (Mussadiq *et al*., 2015). Além da contagem, características como tamanho, forma e coloração podem ser extraídas das imagens e utilizadas na fenotipagem de sementes. Baek *et al.,* (2020) demonstraram a aplicação de análise de imagens para diferentes características de sementes de soja, enquanto Ghimire *et al.,* (2023) obtiveram automaticamente parâmetros como comprimento, largura, área projetada e razão de aspecto a partir de imagens RGB. Informações derivadas da aparência das sementes também têm sido exploradas em sistemas de classificação: Medeiros *et al.,* (2020) utilizaram características visuais de sementes de soja em algoritmos de aprendizado de máquina para classificá-las quanto à aparência e ao potencial fisiológico. Esse contexto torna relevante avaliar não apenas se as sementes são corretamente individualizadas, mas também o comportamento e a interpretação dos descritores obtidos. Vale *et al.,* (2020), em pré-print disponibilizado no arXiv, mostraram que, para sementes de *Fabaceae* sobre fundo azul, a segmentação baseada em contraste de cor separa corretamente objetos sem sobreposição, enquanto Zhao *et al.,* (2023) evidenciaram que a variação de

tamanho, de forma e a sobreposição das sementes afetam o desempenho da contagem automática de soja.

Este trabalho tem como objetivo avaliar a segmentação digital aplicada à contagem automática de sementes de soja de cinco lotes e interpretar os descritores morfométricos e colorimétricos extraídos, comparando a contagem automática à contagem manual de referência e as características das sementes entre os lotes.

Metodologia

O estudo foi conduzido com cinco lotes de sementes de soja (02, 06, 10, 12 e 17), cada um avaliado em quatro repetições de 100 sementes, o que resultou em 20 imagens e 2.000 sementes. Em cada repetição, as sementes foram contadas manualmente, constituindo a contagem manual de referência, e distribuídas em fileiras, sem contato entre si, sobre fundo azul padronizado. As imagens foram adquiridas em escâner HP Scanjet 2004, instalado em posição invertida no interior de uma caixa de alumínio, com resolução de 300 dpi e iluminação própria do equipamento, e armazenadas em formato JPG. Em seguida, cada imagem foi recortada para delimitar a região do fundo azul contendo as sementes. Esse arranjo de aquisição é semelhante ao descrito por Vale *et al.,* (2020).

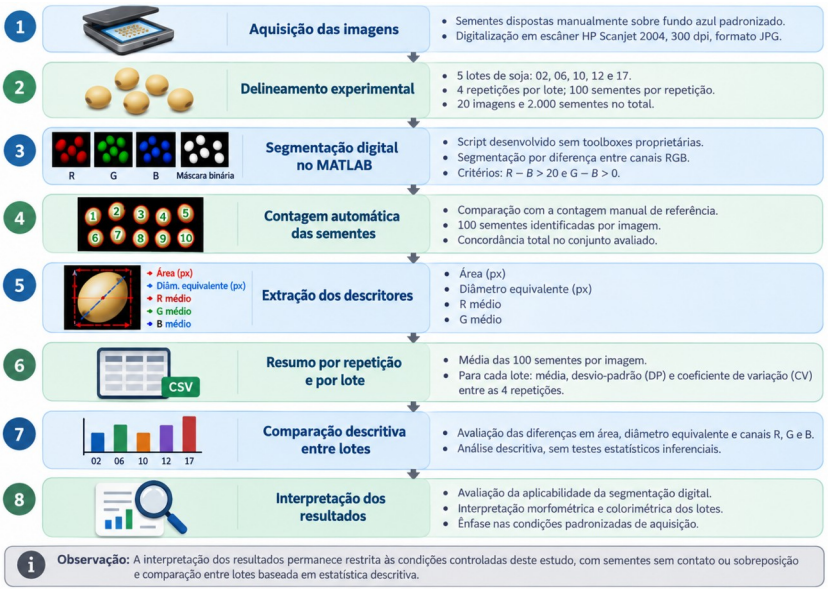
As imagens foram processadas em MATLAB por um script desenvolvido pelo grupo. Na segmentação, foram classificados como sementes os pixels que atenderam simultaneamente às condições $R - B > 20$ e $G - B > 0$, e a máscara binária resultante foi refinada por operações morfológicas de abertura e fechamento. Os objetos foram então individualizados por rotulagem de componentes conexos, e aqueles com área inferior a 500 px foram descartados como ruído. O número de objetos remanescentes em cada imagem foi comparado à contagem manual de referência por meio do erro absoluto e da concordância percentual.

Para cada semente identificada, foram analisados a área (número de pixels do objeto) e o diâmetro equivalente (diâmetro de um círculo de mesma área), como descritores de tamanho, e os valores médios dos canais R, G e B, como descritores de coloração do tegumento.

Os valores individuais foram resumidos pela média das 100 sementes de cada repetição. Para cada lote, calcularam-se a média, o desvio-padrão (DP) e o coeficiente de variação (CV) das quatro repetições, de modo a caracterizar a variabilidade entre repetições e as diferenças descritivas entre lotes. A comparação entre lotes foi realizada de forma descritiva, sem aplicação de testes estatísticos inferenciais.

A Figura 1 resume o fluxo experimental e analítico empregado na avaliação da segmentação e na interpretação dos descritores morfométricos e colorimétricos dos cinco lotes de sementes.

Figura 1 - Esquema metodológico da avaliação da segmentação digital e dos descritores morfométricos e colorimétricos de lotes de sementes de soja.



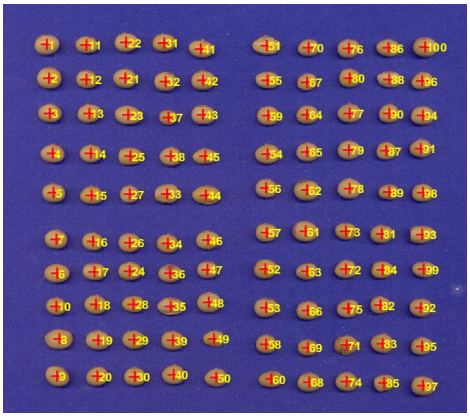
Fonte: elaborado pelos autores (2026).

Resultados

A contagem automática coincidiu com a contagem manual de referência em todas as 20 imagens, com 100 sementes identificadas por imagem e 400 por lote, resultando em erro absoluto nulo e concordância de 100% (2.000/2.000 sementes) (Tabela 1). Nas condições padronizadas adotadas, o contraste entre as sementes e o fundo azul foi suficiente para a individualização dos objetos em todos os lotes.

A Figura 2 apresenta um exemplo do resultado da identificação individual das sementes, correspondente ao Lote 17, repetição R4.

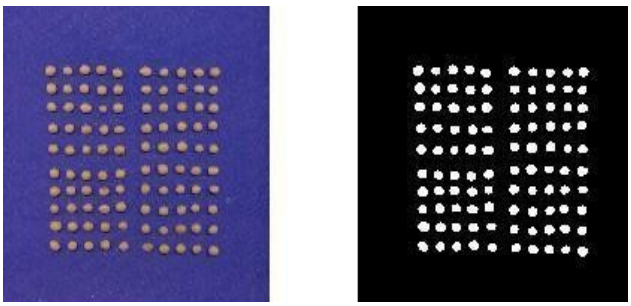
Figura 2 - Exemplo de imagem anotada pelo script, com identificação individual das sementes do Lote 17, repetição R4.



Fonte: elaborado pelos autores (2026).

A Figura 3 ilustra a etapa de segmentação para a mesma repetição, evidenciando a separação das sementes em relação ao fundo azul padronizado.

Figura 3 - Exemplo da segmentação digital do Lote 17, repetição R4: (a) imagem original e (b) máscara binária final.



Fonte: elaborado pelos autores (2026).

Os descritores morfométricos e colorimétricos médios de cada lote são apresentados na Tabela 1. A área média das sementes variou de 4.141 px (Lote 02) a 4.380 px (Lote 10) e o diâmetro equivalente, de 72,5 a 74,5 px, o que corresponde a amplitudes de aproximadamente 6% e 3% entre lotes, respectivamente. O CV da área entre repetições variou de 1,4% (Lote 12) a 10,6% (Lote 17). No Lote 17, esse valor mais elevado esteve associado principalmente à repetição R1 (4.945 px), superior às demais repetições do lote (3.987 a 4.090 px). Nos canais de cor, as médias variaram de 164,2 a 172,7 para R, de 127,7 a 137,5 para G e de 82,4 a 98,6 para B, com CV entre repetições de até 3,0% em

todos os lotes. O canal B apresentou a maior amplitude relativa entre lotes (aproximadamente 20%, contra 5% para R e 8% para G), com o maior valor no Lote 06 e o menor no Lote 12.

Tabela 1 - Contagem de sementes e descritores morfométricos e colorimétricos por lote (média ± DP das 4 repetições, cada uma com 100 sementes).

Lote	Sementes (automática / manual)	Área (px)	Diâmetro equivalente (px)	R	G	B
02	400 / 400	4.141 ± 142	72,5 ± 1,2	165,3 ± 0,8	129,3 ± 1,2	90,9 ± 2,0
06	400 / 400	4.198 ± 261	73,0 ± 2,2	169,3 ± 2,5	135,4 ± 1,5	98,6 ± 0,6
10	400 / 400	4.380 ± 191	74,5 ± 1,6	172,7 ± 2,2	137,5 ± 2,1	93,0 ± 2,8
12	400 / 400	4.322 ± 60	74,1 ± 0,5	164,2 ± 1,3	127,7 ± 0,8	82,4 ± 0,9
17	400 / 400	4.266 ± 454	73,5 ± 3,8	169,9 ± 1,0	133,6 ± 1,2	88,3 ± 1,3

Fonte: elaborado pelos autores (2026).

Discussão

A concordância integral entre a contagem automática e a contagem manual de referência nas 20 imagens indica que a segmentação por contraste de cor foi adequada à condição avaliada, caracterizada por fundo azul homogêneo, sementes espacialmente separadas e ausência de sobreposição, em lotes com diferentes características de tamanho e coloração. Esse resultado é consistente com Vale *et al,* (2020), que também relataram segmentação eficiente de sementes de *Fabaceae* sobre fundo azul contrastante, reforçando a importância da padronização da aquisição para a separação entre objeto e fundo.

Os descritores extraídos ampliam a aplicação do método para além da contagem. O emprego de técnicas de imagem na avaliação de sementes de soja também tem sido explorado como alternativa não destrutiva aos procedimentos convencionais de inspeção, com potencial para a obtenção rápida e objetiva de informações relacionadas à qualidade física das sementes (Mahajan; Mittal; Das, 2018). As diferenças de área e de diâmetro equivalente entre lotes foram da mesma ordem da variação entre repetições de um mesmo lote, especialmente no Lote 17, o que indica que, nas amostras avaliadas, o tamanho médio das sementes pouco diferenciou os lotes. Por outro lado, o canal B apresentou maior amplitude descritiva entre lotes em relação à variabilidade observada entre repetições, indicando que esse parâmetro merece investigação adicional quanto ao seu potencial discriminante. A relevância de atributos de aparência para a diferenciação de sementes também é relatada por Medeiros *et al.,* (2020), que empregaram características derivadas de imagens na classificação de sementes de soja. Entretanto, no presente estudo, a maior amplitude observada no canal B deve ser interpretada apenas como evidência descritiva, uma vez que não foram realizados testes de classificação nem de associação com atributos fisiológicos. Como todas as amostras foram digitalizadas com a iluminação do próprio escâner e sob o mesmo protocolo, essas diferenças tendem a refletir características das sementes, embora a influência de fatores de aquisição não possa ser descartada sem calibração de cor, considerando a sensibilidade de métodos baseados em cor à iluminação (Mussadiq *et al.*, 2015).

A interpretação dos resultados deve, entretanto, permanecer restrita à condição controlada utilizada. Todas as imagens foram adquiridas sob o mesmo protocolo, com sementes distribuídas sem contato ou sobreposição, e cada lote foi representado por 4 repetições, sem aplicação de testes estatísticos de comparação entre lotes. Zhao *et al.,* (2023) apontam que sementes tangentes ou sobrepostas constituem um dos principais desafios para sistemas automáticos de contagem de soja. Assim, os dados obtidos representam evidência de viabilidade do procedimento em condições padronizadas e sustentam seu uso como ferramenta auxiliar para a contagem e a caracterização morfométrica e colorimétrica de lotes de sementes.

Do ponto de vista metodológico, a abordagem empregada corresponde ao processamento convencional de imagens (segmentação por diferença de canais RGB e rotulagem de componentes conexos), e não a métodos de aprendizado profundo (deep learning). Zu *et al.,* (2025) compararam essas duas abordagens na contagem automática de sementes e verificaram que o processamento convencional apresenta acurácia elevada, comparável à contagem manual, quando há condições ambientais controladas, como iluminação uniforme, o que limita sua aplicação em campo. Os métodos

baseados em deep learning, por sua vez, destacaram-se pela velocidade e escalabilidade, mas apresentaram acurácia inconsistente em cenários visualmente complexos ou com sementes densamente agrupadas. Além disso, essas abordagens dependem de bases de imagens anotadas para treinamento e demandam maior complexidade computacional. Os resultados aqui obtidos são coerentes com esse cenário, pois a concordância integral foi alcançada sob fundo padronizado, alto contraste e sementes não sobrepostas. Assim, nenhuma das abordagens pode ser considerada universalmente superior, e a escolha entre elas deve considerar o compromisso entre simplicidade computacional, robustez a variações de cenário e disponibilidade de dados de treinamento.

Conclusão

A segmentação digital apresentou potencial como ferramenta auxiliar para a contagem e caracterização morfométrica e colorimétrica de lotes de sementes de soja em condições padronizadas de aquisição. A ampliação de sua aplicabilidade requer avaliações sob condições menos controladas e análises estatísticas capazes de confirmar as diferenças entre lotes.

Referências

BAEK, J.; LEE, E.; KIM, N.; KIM, S. L.; CHOI, I.; JI, H.; CHUNG, Y. S.; CHOI, M. S.; MOON, J. K.; KIM, K. H. High Throughput Phenotyping for Various Traits on Soybean Seeds Using Image Analysis. **Sensors**, v. 20, n. 1, p. 248, 2020. DOI: 10.3390/s20010248.

GHIMIRE, A.; KIM, S. H.; CHO, A.; JANG, N.; AHN, S.; ISLAM, M. S.; MANSOOR, S.; CHUNG, Y. S.; KIM, Y. Automatic evaluation of soybean seed traits using RGB image data and a Python algorithm. **Plants**, v. 12, n. 17, p. 3078, 2023. DOI: 10.3390/plants12173078.

MAHAJAN, S.; MITTAL, S. K.; DAS, A. Machine vision based alternative testing approach for physical purity, viability and vigour testing of soybean seeds (*Glycine max*). **Journal of Food Science and Technology**, v. 55, n. 10, p. 3949-3959, 2018. DOI: 10.1007/s13197-018-3320-x.

MEDEIROS, A. D.; CAPOBIANGO, N. P.; SILVA, J. M.; SILVA, L. J.; SILVA, C. B.; DIAS, D. C. F. S. Interactive machine learning for soybean seed and seedling quality classification. **Scientific Reports**, v. 10, art. 11267, 2020. DOI: 10.1038/s41598-020-68273-y.

MUSSADIQ, Z.; LASZLO, B.; HELYES, L.; GYURICZA, C. Evaluation and comparison of open source program solutions for automatic seed counting on digital images. **Computers and Electronics in Agriculture**, v. 117, p. 194-199, 2015.

SCHEEREN, B. R.; PESKE, S. T.; SCHUCH, L. O. B.; BARROS, A. C. A. Qualidade fisiológica e produtividade de sementes de soja. **Revista Brasileira de Sementes**, v. 32, n. 3, p. 35-41, 2010.

VALE, A. M. P. G.; UCCHESU, M.; DI RUBERTO, C.; LODDO, A.; SOARES, J. M.; BACCHETTA, G. A new automatic approach to seed image analysis: from acquisition to segmentation. **arXiv preprint**, arXiv:2012.06414, 2020.

ZHAO, J.; KAGA, A.; YAMADA, T.; KOMATSU, K.; HIRATA, K.; KIKUCHI, A.; HIRAFUJI, M.; NINOMIYA, S.; GUO, W. Improved field-based soybean seed counting and localization with feature level considered. **Plant Phenomics**, v. 5, art. 0026, 2023. DOI: 10.34133/plantphenomics.0026.
ZU, Q.; LIU, T.; ZHU, W.; PAN, Y.; WANG, J.; SONG, X.; YU, J.; DANG, S.; YU, X.; ZHANG, Z. Automated seed counting using image processing and deep learning. **Frontiers in Plant Science**, v. 16, art. 1659781, 2025. DOI: 10.3389/fpls.2025.1659781.