

UNIVERSIDADE FEDERAL DE VIÇOSA

Análise de imagens, espectroscopia e aprendizado de máquina na predição do potencial fisiológico e da deterioração de sementes de algodão

Marcelo Augusto Rocha Limão
Doctor Scientiae

VIÇOSA - MINAS GERAIS
2026

MARCELO AUGUSTO ROCHA LIMÃO

Análise de imagens, espectroscopia e aprendizado de máquina na predição do potencial fisiológico e da deterioração de sementes de algodão

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Fitotecnia, para obtenção do título de *Doctor Scientiae*.

Orientadora: Denise C. F. dos S. Dias

Coorientador: Laercio Junio da Silva

**Ficha catalográfica elaborada pela Biblioteca Central da Universidade
Federal de Viçosa - Campus Viçosa**

T

L732a
2026
Limão, Marcelo Augusto Rocha, 1994-
Análise de imagens, espectroscopia e aprendizado de máquina na predição do potencial fisiológico e da deterioração de sementes de algodão / Marcelo Augusto Rocha Limão. – Viçosa, MG, 2026.

1 tese eletrônica (110 f.): il. (algumas color.).

Orientador: Denise Cunha Fernandes dos Santos Dias.
Tese (doutorado) - Universidade Federal de Viçosa,
Departamento de Agronomia, 2026.

Inclui bibliografia.

DOI: <https://doi.org/10.47328/ufvbbt.2026.210>

Modo de acesso: World Wide Web.

1. Algodão - Semente - Fisiologia - Qualidade - Análise.
2. Espectroscopia no infravermelho próximo. 3. Algodão - Semente - Armazenamento - Doenças e danos. 4. Imagens multiespectrais. 5. Raios X. 6. Inteligência artificial. I. Dias, Denise Cunha Fernandes dos Santos, 1958-. II. Universidade Federal de Viçosa. Departamento de Agronomia. Programa de Pós-Graduação em Fitotecnia. III. Título.

CDD 22. ed. 633.5121

MARCELO AUGUSTO ROCHA LIMÃO

Análise de imagens, espectroscopia e aprendizado de máquina na predição do potencial fisiológico e da deterioração de sementes de algodão

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Fitotecnia, para obtenção do título de *Doctor Scientiae*.

APROVADA: 20 de fevereiro de 2026.

Assentimento:

Marcelo Augusto Rocha Limão
Autor

Denise Cunha Fernandes dos Santos Dias
Orientadora

Essa tese foi assinada digitalmente pelo autor em 12/05/2026 às 09:56:08 e pela orientadora em 12/05/2026 às 10:10:14. As assinaturas têm validade legal, conforme o disposto na Medida Provisória 2.200-2/2001 e na Resolução nº 37/2012 do CONARQ. Para conferir a autenticidade, acesse <https://siadoc.ufv.br/validar-documento>. No campo 'Código de registro', informe o código **RFOC.8TUZ.NAZZ** e clique no botão 'Validar documento'.

Aos meus pais (Neide Rocha e Gerismar Limão) cuja dedicação, amor e ensinamentos serviram de pilar para a construção de quem sou.

Dedico

AGRADECIMENTOS

A Deus por toda proteção divina e fé concedida durante essa trajetória, principalmente nos momentos mais difíceis. A TI SENHOR, TODA HONRA E TODA GLÓRIA.

Aos meus pais Neide Rocha e Gerismar Limão por acreditarem no meu potencial e não medirem esforços para sonhar esse sonho comigo.

Aos meus irmãos Mabel Limão e Marcel Limão, meus sobrinhos Ana Cecília, Apollo Limão e Arthur Limão pelo carinho, coragem e incentivo.

À minha vizinha Ana Bessa (in memoriam) por todo amor genuíno e apoio. Sempre foi colo, aconchego e refúgio. Levo-a sempre comigo, no meu coração e nas minhas orações.

À Universidade Federal de Viçosa, ao Departamento de Agronomia e ao Programa de Pós-Graduação em Fitotecnia, pela oportunidade de concluir mais uma etapa fundamental da minha trajetória acadêmica nesta instituição de excelência.

À minha orientadora e professora Denise Dias por todos os ensinamentos, dedicação, paciência, oportunidades, parceria e amizade durante essa trajetória. És um exemplo a ser seguido. Gratidão por tudo.

Ao meu coorientador Laércio Junio pela amizade, ensinamentos e parceria durante essa trajetória.

Aos técnicos do Laboratório de Sementes e integrantes do Grupo de Estudo em Sementes (GSem) pela ajuda, colaboração, parceria e amizade ao longo desses anos.

Aos professores da Universidade Federal de Viçosa e Departamento de Agronomia por todos os ensinamentos e pela significativa contribuição na minha trajetória acadêmica.

Este trabalho foi realizado com o apoio das seguintes agências de pesquisa brasileiras: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) e Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

RESUMO

LIMÃO, Marcelo Augusto Rocha, D.Sc., Universidade Federal de Viçosa, fevereiro de 2026. **Análise de imagens, espectroscopia e aprendizado de máquina na predição do potencial fisiológico e da deterioração de sementes de algodão.** Orientadora: Denise Cunha Fernandes dos Santos Dias. Coorientador: Laercio Junio da Silva.

Sementes de algodão estão sujeitas a danos ao longo do processo produtivo e, por serem oleaginosas, apresentam maior suscetibilidade à deterioração. Devido ao seu alto valor comercial, o controle de qualidade eficiente dos lotes é essencial. Nesse contexto, métodos rápidos, automatizados e, não destrutivos são importantes para complementar os testes tradicionais. O objetivo foi avaliar a eficiência das análises de imagens radiográficas e imagens de plântulas aliadas ao aprendizado interativo de máquina, e das técnicas espectrais para a classificação do potencial fisiológico de sementes de algodão. Buscou-se, ainda, relacionar análises espectrais e alterações bioquímicas e fisiológicas em sementes submetidas à deterioração. Foram utilizadas sementes da cultivar 21065TLP, que foram avaliadas quanto ao potencial fisiológico por meio dos testes de germinação e vigor. Ensaio I - Foi realizado o teste de raios X, sendo as sementes radiografadas e submetidas ao teste de germinação. As imagens foram processadas pelo software ImageJ, obtendo-se as variáveis área, perímetro, preenchimento, densidade relativa e integrada e escala de cinza. A análise de imagens de plântulas foi realizada aos dois, três e quatro dias de germinação, processadas por meio do equipamento GroundEye, obtendo-se as variáveis: comprimento do hipocótilo, da raiz e total das plântulas, índices de crescimento, uniformidade e vigor. Por meio do software Ilastik, foram gerados dados de porcentagem de plântulas normais fortes, normais fracas, anormais e sementes mortas. Ensaio II - Foram realizadas leituras de 200 sementes por lote em espectrofotômetro FT-NIR, na faixa de comprimento de onda de 1.000 a 2.500 nm. Imagens multiespectrais também foram obtidas das mesmas sementes nos comprimentos de 395, 460, 520, 585, 620, 740, 850 e 940 nm, utilizando-se um protótipo equipado com câmera digital e LEDs. Em seguida, estas mesmas sementes foram submetidas ao teste de germinação. Os lotes foram classificados em três categorias de acordo com a germinação: alto (91%), médio (85–90%) e baixo potencial fisiológico (84%). Os dados espectrais NIR originais foram pré-processados utilizando-se os métodos: Standard Normal Variate (SNV); Multiplicative Scatter Correction (MSC); derivadas de Savitzky-Golay (SG) de 1ª e 2ª ordem. Para os dados multiespectrais, foram utilizados SNV e MSC. Os algoritmos de classificação testados foram: Análise Discriminante

por Mínimos Quadrados Parciais (PLS-DA), Random Forest (RF) e Redes Neurais (NN). Ensaio III - Sementes de um lote foram envelhecidas artificialmente por 0 (controle), 24, 48, 72, 96 e 120 h. Foram então realizados testes de germinação, vigor e análises bioquímicas (catalase - CAT, superóxido dismutase – SOD e peroxidase – POX; peróxido de hidrogênio- H₂O₂ e malonaldeído – MDA). Foram também realizadas análises espectrais FT-NIR e multiespectrais, utilizando-se algoritmos de pré-processamento e classificação conforme descrito acima. Foram estabelecidas as classes de germinação: alto (90%), médio (80-89%) e baixo potencial fisiológico (79%). Concluiu-se que, para a técnica de raios X, houve relação entre as variáveis fisiológicas e as variáveis morfobiométricas de densidade relativa, densidade integrada e média de cinza, permitindo identificar sementes bem formadas e a integridade dos tecidos. Aos três dias, variáveis de crescimento de plântulas foram eficientes na classificação dos lotes. O uso do Ilastik apresentou resultados comparáveis aos testes convencionais. Verificou-se, ainda, que a espectroscopia NIR mostrou alta eficiência (até 98% de acurácia), especialmente com PLS-DA. Para imagens multiespectrais, o Random Forest teve melhor desempenho (até 81%). No estudo de deterioração, houve redução da germinação, vigor e atividade enzimática (SOD, CAT, POX), além de aumento de HO e MDA. As técnicas espectrais permitiram classificar sementes em diferentes níveis de deterioração.

Palavras-chave: ft-nir; imagem multiespectral; raios x ; inteligência artificial; qualidade fisiológica.

ABSTRACT

LIMÃO, Marcelo Augusto Rocha, D.Sc., Universidade Federal de Viçosa, February, 2026. **Image analysis, spectroscopy, and machine learning in predicting the physiological potential and deterioration of cotton seeds.** Adviser: Denise Cunha Fernandes dos Santos Dias. Co-adviser: Laercio Junio da Silva.

Cotton seeds are subject to damage throughout the production process and, being oilseeds, exhibit greater susceptibility to deterioration. Due to their high commercial value, efficient quality control of seed lots is essential. In this context, rapid, automated, and non-destructive methods are important to complement traditional tests. The objective was to evaluate the efficiency of radiographic image analysis and seedling images combined with interactive machine learning, and spectral techniques for classifying the physiological potential of cotton seeds. Furthermore, the study sought to relate spectral analyses to biochemical and physiological changes in seeds subjected to restrictions. Seeds of the cultivar 21065TLP were used, and their physiological potential was evaluated through germination and vigor tests. Experiment I – The X-ray test was performed, with the seeds radiographed and subjected to the germination test. The images were processed using ImageJ software, obtaining the variables of area, perimeter, filling, relative and integrated density, and grayscale. Seedling image analysis was performed on days two, three, and four after germination, processed using GroundEye equipment, obtaining the following variables: hypocotyl length, root length, and total seedling length, growth indices, uniformity, and vigor. Using the Ilastik software, data on the percentage of strong normal seedlings, weak normal seedlings, abnormal seedlings, and dead seeds were generated. Experiment II – Readings were taken from 200 seeds per batch using an FT-NIR spectrophotometer, in the wavelength range of 1,000 to 2,500 nm. Multispectral images were also obtained from the same seeds at wavelengths of 395 to 940 nm, using a prototype equipped with a digital camera and LEDs. Subsequently, these same seeds were subjected to the germination test. The seed lots were classified into three categories according to germination: high (91%), medium (85–90%), and low physiological potential (84%). The original NIR spectral data were pre-processed using the following methods: Standard Normal Variate (SNV); Multiplicative Scatter Correction (MSC); and 1st and 2nd order Savitzky-Golay (SG) derivatives. For the multispectral data, SNV and MSC were used. The classification algorithms tested were: Partial Least Squares Discriminant Analysis (PLS-DA), Random Forest (RF), and Neural Networks (NN). Experiment III – Seeds from one lot were subjected to accelerated

aging for 0 (control), 24, 48, 72, 96, and 120 h. Germination, vigor, and biochemical analyses (catalase – CAT, superoxide dismutase – SOD, and peroxidase – POX; hydrogen peroxide – HO and malondialdehyde – MDA) were then performed. FT-NIR and multispectral spectral analyses were also carried out, using the pre-processing and classification algorithms as described above. Germination classes were conditional: high (90%), medium (80-89%), and low physiological potential (79%). It was concluded that, for the X-ray technique, there was a relationship between physiological variables and morphobiometric variables of relative density, integrated density, and grayscale average, allowing the identification of well-formed seeds and tissue integrity. At three days, seedling growth variables were efficient in classifying the lots. The use of Ilastik presented results comparable to conventional tests. It was also found that NIR spectroscopy showed high efficiency (up to 98% accuracy), especially with PLS-DA. For multispectral images, Random Forest performed best (up to 81%). In the interference study, there was a reduction in germination, vigor, and enzymatic activity (SOD, CAT, POX), as well as an increase in HO and MDA. The spectral techniques allowed the seeds to be classified into different amplitude levels.

Keywords: ft-nir; multispectral imaging; x-ray; artificial intelligence; physiological quality.

SUMÁRIO

| | |
|--|----|
| INTRODUÇÃO | 11 |
| REFERÊNCIAS | 16 |
| CAPÍTULO I – Imagens radiográficas e análise de imagens de plântulas para a classificação do potencial fisiológico de sementes de algodão. | 21 |
| RESUMO | 21 |
| ABSTRACT | 23 |
| INTRODUÇÃO | 24 |
| MATERIAL E MÉTODOS | 26 |
| <i>Análises radiográficas</i> | 27 |
| <i>Análise de imagem baseada no desempenho de plântulas pelo GroundEye</i> | 27 |
| <i>Classificação de plântulas por meio do software Ilastik</i> | 28 |
| <i>Delineamento experimental e análise estatística</i> | 30 |
| RESULTADOS E DISCUSSÃO | 31 |
| <i>Caracterização da qualidade fisiológica</i> | 31 |
| <i>Análises radiográficas</i> | 32 |
| <i>Análise de imagem baseada no desempenho de plântulas por meio do software GroundEye</i> | 36 |
| <i>Classificação de plântulas por meio do software Ilastik</i> | 40 |
| CONCLUSÕES | 43 |
| REFERÊNCIAS | 43 |
| CAPÍTULO II – Explorando ferramentas espectrais em conjunto com algoritmos de aprendizado de máquina para classificação do potencial fisiológico de sementes de algodão | 48 |
| RESUMO | 48 |
| ABSTRACT | 49 |
| INTRODUÇÃO | 50 |
| MATERIAIS E MÉTODOS | 52 |
| <i>Espectroscopia no Infravermelho Próximo e Análise de Imagens Multiespectrais</i> | 52 |
| <i>Delineamento Experimental e Análise Estatística</i> | 53 |
| <i>Espectroscopia no Infravermelho Próximo e Análise de Imagens Multiespectrais</i> | 53 |
| RESULTADOS | 54 |
| <i>Grau de umidade e Germinação</i> | 54 |
| <i>Espectroscopia no Infravermelho Próximo (FT-NIR)</i> | 55 |
| <i>Análise de imagens multiespectrais</i> | 58 |
| DISCUSSÃO | 61 |

| | |
|--|-----|
| <i>Caracterização do Potencial Fisiológico</i> | 61 |
| <i>Espectroscopia no infravermelho próximo (NIR)</i> | 62 |
| <i>Análise de imagens multiespectrais</i> | 65 |
| CONCLUSÕES | 67 |
| REFERÊNCIAS | 68 |
| CAPÍTULO III – Alterações bioquímicas, fisiológicas e análises espectrais em sementes de algodão submetidas à deterioração. | 74 |
| RESUMO | 74 |
| ABSTRACT | 76 |
| INTRODUÇÃO | 78 |
| MATERIAL E MÉTODOS | 80 |
| <i>Caracterização do Potencial Fisiológico:</i> | 81 |
| <i>Análises bioquímicas</i> | 81 |
| <i>Espectroscopia no Infravermelho Próximo e Análise de Imagens Multiespectrais</i> | 83 |
| <i>Delineamento Experimental e Análise Estatística</i> | 84 |
| RESULTADOS E DISCUSSÃO | 85 |
| <i>Caracterização do potencial fisiológico</i> | 85 |
| <i>Espectroscopia no infravermelho próximo e análise de imagens multiespectrais</i> | 91 |
| CONCLUSÕES | 99 |
| CONCLUSÕES GERAIS | 99 |
| REFERÊNCIAS | 100 |

INTRODUÇÃO

O algodão (*Gossypium hirsutum* L.) é a principal fibra natural de importância econômica em escala global, figurando como uma das mais relevantes culturas industriais. No contexto brasileiro, a cotonicultura destaca-se como uma das principais *commodities* agrícolas, em função do elevado nível de tecnificação dos sistemas produtivos e das condições edafoclimáticas conforme as exigências da cultura (Baio et al., 2020). Na safra 2024/2025, a área de algodão cultivada no Brasil alcançou aproximadamente 2,14 milhões de hectares, com destaque para os estados de Mato Grosso e Bahia, que concentram a maior parte da produção e lideram a expansão da área plantada (ABRAPA, 2025a). No mesmo período, a cadeia produtiva do algodão movimentou cerca de US\$ 4.813.536,119 na balança comercial brasileira, registrando recordes históricos de exportação e evidenciando a relevância econômica do setor para o agronegócio nacional (ABRAPA, 2025b).

Em sistemas produtivos de alta tecnologia, como os praticados no cultivo de algodão no Brasil, a garantia de uma colheita bem-sucedida e rentável começa pela qualidade do insumo inicial: a semente (Dowd et al., 2019). Sementes de algodão estão sujeitas a uma série de danos como os causados por deterioração em campo, pragas e doenças na fase de pré-colheita, ou danos mecânicos durante as etapas de colheita, descaroçamento, deslintamento e processamento. Além disto, por serem oleaginosas, são mais propensas à deterioração, principalmente durante o armazenamento (Shahein et al., 2022). Esses aspectos podem afetar negativamente a sua qualidade fisiológica, o estabelecimento das plântulas e a uniformidade do estande em campo.

Considerando que são sementes de alto valor comercial, as expectativas são sempre relacionadas à obtenção de um alto desempenho em campo após a semeadura. Portanto, obter informações seguras sobre a qualidade das sementes durante as etapas mencionadas acima é fundamental para um controle de qualidade eficiente dos lotes produzidos a cada safra, propiciando posicionar neste mercado de alto valor os lotes de sementes com maior potencial fisiológico.

Tradicionalmente, para a aferição e controle de qualidade das sementes, são realizados testes de germinação (Brasil, 2025) e de vigor (Krzyzanowski et al., 2020). Embora eficientes, são destrutivos, gerando resíduos e requerem pelo menos 12 dias para a obtenção dos resultados, dificultando agilizar o controle de qualidade realizado pelas empresas de sementes nas etapas de beneficiamento e armazenamento. Este cenário tem impulsionado a busca por métodos

rápidos, não destrutivos e automatizados que possam ser integrados aos programas de controle de qualidade (Gomes-Junior, 2020).

Dentre as tecnologias emergentes, podemos destacar a análise de imagem de raios X (França-Silva et al., 2023), análise de imagens baseada no desempenho de plântulas pelo *software* GroundEye (Manual GroundEye®, 2016), análise de imagens baseada em aprendizado de máquina interativo (IML) (Berg, 2019; Limão et al., 2023), a espectroscopia no infravermelho próximo (Xia et al., 2019) e a análise de imagens multiespectrais (Petronilio et al., 2025).

A análise de raios X baseia-se no princípio de que a radiação eletromagnética é absorvida de forma diferenciada pelas estruturas internas da semente, dependendo de sua densidade (Simak e Gustafsson, 1953; ISTA, 2020). Em sementes, essa ferramenta pode ser utilizada para a inspeção rápida de características internas, permitindo a identificação objetiva e a quantificação de danos internos, como o grau de preenchimento do embrião, a presença de danos mecânicos e a ocorrência de infestação por insetos (França-Silva et al., 2023; Nogueira et al., 2024). Ao fornecer informações sobre a integridade física, a análise radiográfica pode permitir estabelecer correlação com o potencial fisiológico das sementes, como observado em estudos com sementes de soja (França-Silva et al., 2023; Wendt et al., 2014), mamão (Limão et al., 2025; Nogueira et al., 2024), crambe (Ribeiro et al., 2021), leucena (Medeiros et al., 2018) e brócolis (Abud et al., 2018). Em sementes de algodão esta técnica também se mostrou promissora, conforme o estudo de Sawarkar e Rode (2017).

Já o GroundEye é um sistema avançado de análise de imagens, constituído por um *software* integrado a uma câmera digital de alta resolução e sistemas de iluminação que pode ser utilizado para automatizar avaliações de sementes e de plântulas (Brandani et al, 2023; Manual Groundeye®, 2016). Sua principal aplicação reside na quantificação objetiva de parâmetros morfométricos das plântulas, como o comprimento radicular, do hipocótilo, comprimento de plântulas e cálculo dos índices de vigor, crescimento e uniformidade (Brandani et al, 2023; Manual Groundeye®, 2016). O GroundEye utiliza algoritmos que automatizam a segmentação e a medição, permitindo a classificação rápida de plântulas quanto ao vigor, fornecendo dados detalhados que podem se correlacionar com o potencial de estabelecimento da cultura no campo, superando a subjetividade e a morosidade das medições manuais tradicionais (Brandani et al, 2023; Manual Groundeye®, 2016). Embora não haja relatos do uso dessa técnica para avaliação do potencial fisiológico das sementes de algodão, foi utilizada para soja (Acha e Vieira, 2020; Pinto et al., 2018), milho (Catão et al., 2020; Ferreira et al., 2020),

café (Abreu et al., 2016), repolho e salsa (Vasconcelos et al., 2018). O sistema GroundEye também foi eficiente na caracterização morfológica e fisiológica de sementes de *Amaranthus* spp., permitindo separar as três espécies do gênero em estudo (Xavier et al., 2019). Possibilitou também a caracterização de sementes e embriões de *Strelitzia reginae* (Taniguchi et al., 2022).

Baseado em IML, o Ilastik é um *software* de código aberto que se destaca por sua interface didática e interativa, sendo cada vez mais empregado na análise e classificação de sementes e plântulas (Medeiros et al., 2020a; Limão et al., 2023). Sua importância reside na capacidade de permitir que os usuários treinem modelos de classificação baseados em características fornecidas por eles, eliminando a necessidade de programação complexa (Yasam et al., 2022). Especificamente no controle de qualidade de sementes, o Ilastik facilita a segmentação precisa de diferentes características tanto de sementes (distinção de coloração de tegumentos, estrutura de sementes etc.) como de plântulas (plântulas normais, anormais, fortes e fracas etc.) (Lin et al., 2019; Medeiros et al., 2020a), automatizando a classificação de parâmetros morfológicos e estruturais complexos com alta eficiência e repetibilidade, essenciais para o controle de qualidade automatizado e para a agilidade nas tomadas de decisão neste processo (Dietz et al., 2020). Limão et al. (2023) realizaram a classificação do vigor de sementes de lentilha com base em técnicas de análise de imagens de plântulas e aprendizado de máquina interativo e concluíram que os dados obtidos pelo Ilastik, após 4 dias de germinação, permitiram o desenvolvimento de modelos com alta precisão para avaliação do vigor das sementes. Outros estudos comprovaram a eficiência deste *software* em sementes de soja (Medeiros et al., 2020a), arroz (Yasam et al., 2022), braquiária (Souza et al., 2023) e pinhão-mansão (Medeiros et al., 2020b).

Na atualidade, a aplicação de tecnologias espectrais, como a espectroscopia no infravermelho próximo (Xia et al., 2019) e análise de imagens multiespectrais (Petronilio et al., 2025), combinadas com algoritmos de aprendizagem de máquina também tem demonstrado potencialidade na avaliação do potencial fisiológico de sementes, como métodos complementares aos tradicionais e superando as limitações destes (Xia et al., 2019). Tais ferramentas baseiam-se na interação das sementes com a radiação eletromagnética em diferentes faixas espectrais, que pode variar conforme a sua composição química, ou seja, pela presença de água e de compostos orgânicos como carboidratos, lipídios, proteínas e outros (Xia et al., 2019). É possível, então, associar alterações nestes compostos relacionadas ao nível de deterioração das sementes, e que afetam o seu potencial fisiológico (Xia et al., 2019).

Em nível celular, o processo de deterioração das sementes resulta do desequilíbrio entre a produção e a eliminação de espécies reativas de oxigênio (EROs), causando estresse oxidativo. EROs, como o radical superóxido ($O_2^{\bullet-}$), o peróxido de hidrogênio (H_2O_2) e o radical hidroxila (OH^{\bullet}), são geradas continuamente como subprodutos do metabolismo respiratório mitocondrial, da peroxidação de lipídios e de diversas reações enzimáticas, promovendo danos celulares quando acumuladas em níveis elevados (Pinheiro et al., 2023; Krzyzanowski; Dias; França-Neto, 2022). A magnitude desses danos está diretamente relacionada à capacidade das células em remover ou neutralizar as EROs por meio de sistemas antioxidantes eficientes. Nesse contexto, o sistema antioxidante enzimático, constituído pelas enzimas superóxido dismutase (SOD), catalase (CAT), peroxidase do ascorbato (APX) e peroxidases (POX), dentre outras, desempenha papel essencial na manutenção da homeostase celular (Mittler, 2017). A redução da atividade destas enzimas pode intensificar o estresse oxidativo, acelerando o processo de deterioração e o declínio do vigor das sementes, conforme observado por Xing et al. (2023).

Portanto, considerando que a deterioração é um processo oxidativo que promove alterações nos constituintes químicos das sementes, devido principalmente à peroxidação lipídica, técnicas de espectroscopia e de análises multiespectrais podem ser promissoras para classificar lotes de sementes quanto à qualidade fisiológica, de acordo com o seu nível e deterioração (Xia et al., 2019). Tais técnicas são interessantes para serem testadas em sementes de algodão, que por serem oleaginosas são mais propensas à deterioração, devido à peroxidação lipídica (Shahein et al., 2022).

A espectroscopia no infravermelho próximo (FT-NIR) baseia-se na absorção pelas sementes de radiação eletromagnética nos comprimentos de onda entre 780-2500 nm (Xia et al., 2019). A interação das sementes com a radiação eletromagnética que é absorvida pela água e compostos orgânicos como carboidratos, lipídios e proteínas permite associar alterações na sua composição química, relacionadas ao processo de deterioração, que afetam o seu potencial fisiológico (Xia et al., 2019). Para isto, há necessidade de se utilizar algoritmos de pré-processamento como o *Multiplicative Scatter Correction* (MSC), o *Standard Normal Variate* (SNV) e as derivadas de 1ª e 2ª ordem de Savitsky e Golay, com o objetivo de corrigir os efeitos aditivos e multiplicativos ocasionados pela amostra e/ou pelo equipamento, e assim obter uma assinatura espectral limpa (Delwiche, 1998; Savitsky; Golay, 1964). O desenvolvimento dos modelos de classificação dos lotes de sementes quanto ao potencial fisiológico tem como base os algoritmos de aprendizado de máquina, pois podem capturar relações lineares e não lineares com alta precisão (Medeiros et al., 2020b). Assim, os espectros brutos são processados por meio

de algoritmos como a Regressão por Mínimos Quadrados Parciais – PLS, *Random Forest* ou Redes Neurais para criar modelos preditivos que correlacionam a assinatura espectral com atributos de qualidade das sementes (Medeiros et al., 2020b; Burns e Ciureczak, 2007).

Resultados interessantes na avaliação da qualidade de sementes vêm sendo obtidos para diversas espécies como soja (Soares et al., 2024), milho (Andriazzi, Rocha e Custódio, 2023), braquiária (Souza et al., 2023), milho doce (Qiu et al., 2018, Queiróz et al., 2025). No algodoeiro, a técnica foi utilizada para distinguir diferentes genótipos (Mata et al., 2022) e na avaliação da qualidade fisiológica de sementes com e sem línter (Mayrinck et al., 2020).

Em relação às imagens multiespectrais, essa tecnologia possibilita inferir sobre características físicas e químicas das sementes, fornecendo dados espaciais e espectrais utilizando comprimentos de onda específicos que permitem avaliar atributos de qualidade como estrutura de superfície, textura, composição química e, potencialmente, aspectos fisiológicos (Medeiros et al., 2020a). Assim, a luz, ao interagir com a matéria, é absorvida, refletida ou transmitida em comprimentos de onda específicos, dependendo da composição química da amostra (Galletti et al., 2020; Petronilio et al., 2025). Em sementes, as imagens multiespectrais criam uma "assinatura química", permitindo extrair marcadores ópticos referentes à composição química, conteúdo de fluoróforos como clorofila *a*, clorofila *b*, lignina e atributos físicos das sementes como cor, brilho, textura e área (França-Silva et al., 2023). Essas informações espectrais predizer o potencial fisiológico das sementes, permitindo inferir de forma não destrutiva, precisa e rápida (Petronilio et al., 2025).

O uso dessa técnica foi comprovado para a identificação de sementes de cultivares de tomate (Shrestha et al., 2015) e trigo (Vresak et al., 2016), na classificação de diferentes níveis de vigor em sementes de soja utilizando autofluorescência (Silva et al., 2021), no potencial fisiológico de sementes de cenoura e tomate avaliado pela reflectância e fluorescência de clorofila (Galletti et al., 2020) e na identificação de sementes de soja produzidas sob condições favoráveis e de estresse por meio da fluorescência de clorofila (Petronilio et al., 2025).

Em síntese, a integração das tecnologias de espectroscopia no infravermelho próximo e de análises por imagens multiespectrais pode ser promissora para a avaliação do potencial fisiológico das sementes de algodão de forma automatizada, rápida, assertiva e não destrutiva. Tais ferramentas têm potencial para a utilização nos programas de controle de qualidade, especialmente durante as etapas de pós-colheita agilizando a identificação dos lotes de alto, médio e baixo potencial fisiológico produzidos a cada safra.

REFERÊNCIAS

ABRAPA. Análises históricas de área, produção e produtividade do algodão no Brasil. 2025a. Disponível

em:<https://app.powerbi.com/view?r=eyJrIjoiNDZhMDIhZjYtZTljNi00NmLWIwMDEtOGY4OGE1OTE5ZmE1IiwidCI6IjRhMDk1OGIzLTg4MWQtNDBmYS05NTU1LTlwODQ1MzdhYWZkMyJ9&pageName=ReportSection10c4aa5d90c47641852d>. Acesso em: 5 nov. 2025.

ABRAPA. Evolução da balança comercial anual de algodão do Brasil e UFs. 2025b. Disponível em:<https://app.powerbi.com/view?r=eyJrIjoiMzlmMzNiMjltMjg4Mi00NWVhYzQtMmVlMmI4OGQ5YjU0IiwidCI6IjRhMDk1OGIzLTg4MWQtNDBmYS05NTU1LTlwODQ1MzdhYWZkMyJ9>. Acesso em: 5 nov. 2025.

ABREU, L. A. S. et al. Computerized analysis in the physiological quality of coffee seeds. *International Journal of Current Research*, v. 8, n. 11, p. 40820-40823, 2016.

ABUD, H. F. et al. Radiographic images and relationship of the internal morphology and physiological of broccoli seed. *Acta Scientiarum: Agronomy*, v. 40, n. 1, p. 1-9, 2018.

ACHA, A. J.; VIEIRA, H. D. Digital image processing of coated perennial-soybean seeds and correlation with physiological attributes. *Journal of Seed Science*, v. 42, e202042004, 2020.

ANDRIAZZI, C. V. G.; ROCHA, D. K.; CUSTÓDIO, C. C. Determination of the physiological quality of corn seeds by infrared equipment. *Journal of Seed Science*, v. 45, 2023. DOI: <https://doi.org/10.1590/2317-1545v45265346>.

BAIO, F. H. R. et al. Application technology of boron via foliar and its effects on cotton crop phenology. *Brazilian Journal of Development*, v. 2, n. 6, p. 7367-7379, 2020.

BERG, S. et al. Ilastik: interactive machine learning for (bio) image analysis. *Nature Methods*, v. 16, n. 12, p. 1226-1232, 2019. DOI: <https://doi.org/10.1038/s41592-019-0582-9>.

BRANDANI, E. B. et al. Image analysis for the evaluation of soybean seeds vigor. *Acta Agronômica*, v. 70, n. 3, p. 311-316, 2021. DOI: <https://doi.org/10.15446/acag.v70n3.92093>.

BRASIL. Ministério da Agricultura e Pecuária. Secretaria de Defesa Agropecuária. Regras para análise de sementes – RAS. Brasília: MAPA/SDA, 2025. Disponível em: https://wikisda.agricultura.gov.br/pt-br/Laborat%C3%B3rios/Metodologia/Sementes/RAS_2025/cap_15_tol_fiscalizacao.

BURNS, D. A.; CIURCZAK, E. W. *Handbook of near-infrared analysis*. 3. ed. Boca Raton: CRC Press, 2007.

CATÃO, H. C. R. M. et al. Antioxidant activity and physiological performance of popcorn seed after saline stress and analysis of seedling images. *Ciência e Agrotecnologia*, v. 44, e005020, 2020. DOI: <https://doi.org/10.1590/1413-7054202044005020>.

DELWICHE, S. R. Protein content of single kernels of wheat by near-infrared reflectance spectroscopy. *Journal of Cereal Science*, v. 27, n. 3, p. 241-254, 1998.

DIETZ, C. et al. Integration of the ImageJ ecosystem in the KNIME analytics platform. *Frontiers in Computer Science*, v. 2, 2020. DOI: <https://doi.org/10.3389/fcomp.2020.00008>.

DOWD, M. K.; MANANDHAR, R.; DELHOM, C. D. Effect of seed orientation, acid delinting, moisture level, and sample type on cottonseed fracture resistance. *Transactions of the ASABE*, v. 62, n. 4, p. 1045-1053, 2019.

FERREIRA, O. J. M. et al. Tecnologia de análise de imagens para a seleção de sementes crioulas de milho. *Global Science and Technology*, v. 13, n. 2, p. 28-38, 2020.

FRANÇA-SILVA, F. et al. Advances in imaging technologies for soybean seed analysis. *Journal of Seed Science*, v. 45, e202345022, 2023. DOI: <https://doi.org/10.1590/2317-1545v45274098>.

GALLETTI, P. A. et al. Integrating optical imaging tools for rapid and non-invasive characterization of seed quality. *Frontiers in Plant Science*, v. 11, 577851, 2020. DOI: <https://doi.org/10.3389/fpls.2020.577851>.

GOMES-JUNIOR, F. G. Análise computadorizada de imagens de plântulas. In: KRZYZANOWSKI, F. C. et al. (ed.). *Vigor de sementes: conceitos e testes*. Londrina: ABRATES, 2020. p. 142-176.

INTERNATIONAL SEED TESTING ASSOCIATION. *International rules for seed testing*. Bassersdorf: ISTA, 2020.

KRZYZANOWSKI, F. C. et al. (ed.). *Vigor de sementes: conceitos e testes*. Londrina: ABRATES, 2020.

KRZYZANOWSKI, F. C.; DIAS, D. C. F. S.; FRANÇA-NETO, J. F. B. Deterioração e vigor da semente. Londrina: EMBRAPA, 2022.

- LIMÃO, M. A. R. et al. Classification of lentil seed vigor based on seedling image analysis techniques and interactive machine learning. *Journal of Seed Science*, v. 45, e202345038, 2023.
- LIMÃO, M. A. R. et al. X-ray and tetrazolium tests for assessing papaya seed viability. *Semina: Ciências Agrárias*, v. 46, n. 5, p. 1593-1610, 2025. DOI: <https://doi.org/10.5433/1679-0359.2025v46n5p1593>.
- LIN, P. et al. Determinação rápida e exata da qualidade da semente de soja seca pós-colheita com base na tecnologia de visão mecânica. *Scientific Reports*, v. 9, p. 1-11, 2019.
- MANUAL GROUNDEYE®. Lavras: Tbit Tecnologia e Sistemas, 2016.
- MATA, M. M. et al. Distinguishing cotton seed genotypes by means of vibrational spectroscopic methods (NIR and Raman) and chemometrics. *Spectrochimica Acta Part A*, v. 266, 120399, 2022. DOI: <https://doi.org/10.1016/j.saa.2021.120399>.
- MAYRINCK, L. G. et al. Use of near infrared spectroscopy in cotton seeds physiological quality evaluation. *Journal of Seed Science*, v. 42, p. 1-11, 2020. DOI: <https://doi.org/10.1590/2317-1545v42227169>.
- MEDEIROS, A. D. et al. Interactive machine learning for soybean seed and seedling quality classification. *Scientific Reports*, v. 10, 11267, 2020a. DOI: <https://doi.org/10.1038/s41598-020-68273-y>.
- MEDEIROS, A. D. et al. Classificação de qualidade de sementes de pinhão-manso utilizando imagens radiográficas e aprendizado de máquina. *Industrial Crops and Products*, v. 146, e112162, 2020b. DOI: <https://doi.org/10.1016/j.indcrop.2020.112162>.
- MEDEIROS, A. D. et al. Parameters based on X-ray images to assess the physical and physiological quality of *Leucaena leucocephala* seeds. *Ciência e Agrotecnologia*, v. 42, n. 6, p. 643-652, 2018. DOI: <https://doi.org/10.1590/1413-70542018426023318>.
- MITTLER, R. ROS are good. *Trends in Plant Science*, v. 22, n. 1, p. 11-19, 2017.
- NOGUEIRA, G. A. et al. X-rays for determination of papaya seeds viability. *Revista Ciência Agronômica*, v. 55, e20238712, 2024. DOI: <https://doi.org/10.5935/1806-6690.20240021>.
- PETRONILIO, A. C. P. et al. Smart and accurate: a new tool to identify stressed soybean seeds based on multispectral images and machine learning models. *Smart Agricultural Technology*, v. 12, 101042, 2025.

- PINHEIRO, D. T. et al. Oxidative stress, protein metabolism, and physiological potential of soybean seeds under weathering deterioration. *Acta Scientiarum: Agronomy*, v. 45, e56910, 2023. DOI: <https://doi.org/10.4025/actasciagron.v45i1.56910>.
- PINTO, C. A. G. et al. Relationship between size and physiological potential of soybean seeds. *Seed Science and Technology*, v. 46, n. 3, p. 497-510, 2018.
- QIU, G. et al. Single-kernel FT-NIR spectroscopy for detecting supersweet corn seed viability. *Sensors*, v. 18, n. 4, 1010, 2018. DOI: <https://doi.org/10.3390/s18041010>.
- QUEIROZ, G. B. et al. Computerized analysis of seedlings and near-infrared spectroscopy in the evaluation of sweet corn seeds. *Journal of Seed Science*, v. 47, e202547032, 2025. DOI: <https://doi.org/10.1590/2317-1545v47302324>.
- RIBEIRO, A. S. et al. Use of radiographic images for assessment of crambe seed quality. *Journal of Seed Science*, v. 43, e202143001, 2021.
- SAVITZKY, A.; GOLAY, M. J. E. Smoothing and differentiation of data by least squares procedures. *Analytical Chemistry*, v. 36, n. 8, p. 1627-1639, 1964.
- SAWARKAR, M. R.; RODE, S. V. X-ray based germination test of cotton seed. *International Journal of Advanced Research in Computer Science*, v. 8, n. 5, 2017. DOI: <https://doi.org/10.26483/ijarcs.v8i5.3429>.
- SHAHEIN, A. M. E. et al. Effect of storage period and condition on cotton seed viability. *Journal of the Advances in Agricultural Researches*, v. 27, n. 3, p. 582-591, 2022. DOI: <https://doi.org/10.21608/jalexu.2022.161397.1080>.
- SHRESTHA, S. et al. Use of multispectral imaging in varietal identification of tomato. *Sensors*, v. 15, p. 4496-4512, 2015.
- SILVA, C. B. et al. Autofluorescence-spectral imaging for soybean seed quality. *Scientific Reports*, v. 11, 17834, 2021. DOI: <https://doi.org/10.1038/s41598-021-97223-5>.
- SIMAK, M.; GUSTAFSSON, Å. X-ray photography in seed tests. *Hereditas*, v. 39, p. 458-468, 1953.
- SOARES, J. M. et al. Classification of soybean seed lots using infrared spectroscopy. *Journal of Seed Science*, v. 46, 2024.

- SOUZA, L. R. et al. Near infrared spectroscopy and seedling image analysis in *Urochloa decumbens*. *Journal of Seed Science*, v. 45, 2023. DOI: <https://doi.org/10.1590/2317-1545v4527702>.
- TANIGUCHI, M. et al. Morpho-anatomical characterization of *Strelitzia reginae*. *Ornamental Horticulture*, v. 28, n. 3, p. 314-322, 2022.
- VASCONCELOS, M. C. et al. Radiography and biometric analysis of vegetable seeds. *Amazonian Journal of Agricultural and Environmental Sciences*, v. 61, 2018.
- VRESAK, M. et al. Image-spectroscopy technology for seed testing. *PLoS One*, v. 11, e0152011, 2016. DOI: <https://doi.org/10.1371/journal.pone.0152011>.
- WENDT, L. et al. Avaliação do potencial fisiológico de sementes de soja. *Pesquisa Agropecuária Tropical*, v. 44, n. 3, p. 280-286, 2014.
- XAVIER, J. B. et al. Characterization of *Amaranthus spp.* seeds. *Journal of Seed Science*, v. 41, p. 478-487, 2019. DOI: <https://doi.org/10.1590/2317-1545v41n4226286>.
- XIA, Y. et al. Non-destructive detection of seed viability. *Artificial Intelligence in Agriculture*, v. 1, p. 35-47, 2019.
- XING, M. et al. Physiological alterations and nondestructive test methods of crop seed vigor. *Agriculture*, v. 13, n. 3, 527, 2023. DOI: <https://doi.org/10.3390/agriculture13030527>.
- YASAM, S. et al. Previsão da germinação baseada em aprendizagem supervisionada. *Soft Computing*, v. 26, p. 13133-13144, 2022. DOI: <https://doi.org/10.1007/s00500-022-06910-6>

CAPÍTULO I – Imagens radiográficas e análise de imagens de plântulas para a classificação do potencial fisiológico de sementes de algodão.

RESUMO

A demanda por métodos rápidos e precisos está cada vez mais crescente na tecnologia de sementes, o que torna o teste de raios X, a análise computadorizada de plântulas pelo GroundEye e o Ilastik alternativas promissoras para avaliação do potencial fisiológico das sementes. O objetivo do trabalho foi avaliar a eficiência de imagens radiográficas e de imagens plântulas aliadas ao aprendizado interativo de máquina para a classificação do potencial fisiológico de sementes de algodão. Sementes de oito lotes da cv. 21065TLP foram caracterizadas quanto à qualidade fisiológica pelos testes de germinação e vigor. Para aquisição das imagens de raios X, 200 sementes de cada lote foram identificadas e radiografadas. As mesmas sementes foram submetidas ao teste de germinação. As imagens radiográficas obtidas foram processadas e avaliadas por meio do software ImageJ, obtendo-se as seguintes variáveis: área, perímetro, preenchimento, densidade relativa, densidade integrada e escala de cinza. Para a análise computadorizada de imagens de plântulas, as sementes foram colocadas para germinar a 25 °C por dois, três e quatro dias. As imagens foram capturadas e processadas em equipamento GroundEye, obtendo-se as variáveis: comprimento do hipocótilo, da raiz e total das plântulas, índices de crescimento, de uniformidade e de vigor. Por meio do software Ilastik, foram gerados dados e convertidos em porcentagem de se obter plântulas normais fortes, normais fracas, anormais e sementes mortas. Foi utilizado o delineamento inteiramente casualizado, com quatro repetições. Os dados foram submetidos à análise de variância e as médias obtidas para cada lote foram comparadas pelo teste de Tukey ($p < 0,05$). Também foi realizada a análise de componentes principais (PCA) utilizando-se os dados obtidos nos testes de caracterização do potencial fisiológico e as variáveis obtidas por meio das análises de imagens radiográficas das sementes e de crescimento das plântulas. A técnica de raios X apresenta relação com o potencial fisiológico das sementes de algodão, além de fornecer informações sobre suas estruturas internas. Aos três dias de germinação, o comprimento da parte aérea, da raiz primária, o comprimento total da plântula e os índices de uniformidade, crescimento e vigor, determinados por meio do GroundEye, são eficientes na classificação de lotes de algodão quanto ao potencial fisiológico. O aprendizado de máquina supervisionado por meio do Ilastik possibilitou a classificação dos lotes quanto ao potencial fisiológico, obtendo resultados comparáveis aos testes convencionais.

Palavras-chave: qualidade fisiológica, *machine learning*, GroundEye, Ilastik, *Gossypium hirsutum*.

ABSTRACT

The demand for rapid and precise methods is increasingly growing in seed technology, making the X-ray test, computerized seedling analysis using GroundEye, and Ilastik promising alternatives for evaluating seed physiological potential. The objective of this study was to evaluate the efficiency of radiographic images and seedling images combined with interactive machine learning for classifying the physiological potential of cotton seeds. Seeds from eight lots of the cultivar 21065TLP were characterized for physiological quality using germination and vigor tests. For X-ray image acquisition, 200 seeds from each lot were identified and radiographed. The same seeds were subsequently subjected to the germination test. The obtained radiographic images were processed and analyzed using the ImageJ software, and the following variables were measured: area, perimeter, filling, relative density, integrated density, and gray scale. For computerized seedling image analysis, seeds were germinated at 25 °C for two, three, and four days. Images were captured and processed using the GroundEye system, obtaining the following variables: hypocotyl length, root length, total seedling length, and growth, uniformity, and vigor indices. By Ilastik software, data were generated and converted into percentages of strong normal seedlings, weak normal seedlings, abnormal seedlings, and dead seeds. A completely randomized design with four replications was used. Data were subjected to analysis of variance, and the means for each lot were compared using Tukey's test ($p < 0.05$). Principal component analysis (PCA) was also performed using data obtained from physiological potential tests, as well as variables derived from radiographic image analysis of seeds and seedling growth. The X-ray technique showed a relationship with the physiological potential of cotton seeds, in addition to providing information about their internal structures. At three days of germination, shoot length, primary root length, total seedling length, and the uniformity, growth, and vigor indices, as determined by GroundEye, were efficient for classifying cotton seed lots according to physiological potential. Supervised machine learning using Ilastik also enabled the classification of seed lots according to physiological potential, with results comparable to conventional tests.

Keywords: physiological quality, machine learning, GroundEye, Ilastik, *Gossypium hirsutum*.

INTRODUÇÃO

O algodoeiro (*Gossypium hirsutum* L.) é a principal fonte de fibra comercial na indústria têxtil no mundo, sendo o Brasil um dos cinco maiores produtores e o maior exportador mundial na safra de 2024/2025, superando os Estados Unidos (ABRAPA, 2025). Considerando que o processo de produção de algodão é altamente tecnificado, a utilização de sementes de alta qualidade é fundamental para o sucesso das lavouras, garantindo estandes adequados, plantas uniformes com reflexos positivos na produtividade (Rehman et al., 2020).

Rotineiramente, o potencial fisiológico das sementes é avaliado pelos testes tradicionais de germinação e vigor; entretanto, estes testes são destrutivos, demorados e envolvem interpretações humanas, que podem ser subjetivas, tornando-se um fator limitante para um controle de qualidade mais ágil (Elmasry et al., 2019; Xia et al., 2019). Neste contexto, a aplicação de tecnologias modernas baseada em análise de imagem e aprendizado interativo de máquinas tornam-se alternativas potenciais para a avaliação da qualidade de sementes, superando as limitações enfrentadas pelos testes tradicionais (Xia et al., 2019; Limão et al., 2023). Dentre estas técnicas, destacam-se a análise de raios X e a análise computadorizada de imagens de plântulas, que possuem vantagens como permitir avaliação rápida das estruturas internas das sementes e do desempenho de plântulas, fornecer resultados confiáveis e um banco de dados e imagens que podem ser armazenados e referenciados conforme necessário (Gomes-Junior, 2020).

O teste de raios X baseia-se na interação das ondas eletromagnéticas com a semente, com comprimento de onda que pode variar de 0,01 a 10 nanômetros e energia de 0,1 a 100 keV (Kotwaliwale et al., 2014). A penetração dos raios X é influenciada pela estrutura, composição e densidade dos tecidos da semente, isto é, tecidos densos absorvem mais radiação, resultando em áreas claras, enquanto tecidos menos densos absorvem menos radiação, resultando em áreas mais escuras na imagem radiográfica (Simak e Gustafsson, 1953; ISTA, 2004). Esta técnica permite verificar de forma não destrutiva alterações na densidade da amostra que podem ser relacionadas à sua integridade, comprovado em diferentes estudos para sementes de soja (França-Silva et al., 2023; Wendt et al., 2014), mamão (Nogueira et al., 2024), crame (Ribeiro et al., 2023), leucena (Medeiros et al., 2018) e brócolis (Abud et al., 2018).

Nos últimos anos, a análise computadorizada de imagens tem sido amplamente utilizada na tecnologia de sementes para avaliar a qualidade fisiológica, fundamentando-se principalmente em características morfológicas de plântulas. A captura das imagens ocorre de forma rápida, por meio de câmeras digitais, escâner ou até mesmo celulares, e sua interpretação

é realizada por algoritmos específicos capazes de identificar atributos de interesse e comparar diferentes lotes (Gomes-Junior, 2020). Entre as vantagens desse método, destacam-se a maior precisão e rapidez das avaliações, que podem ser concluídas em poucos segundos; a possibilidade de padronização entre laboratórios e analistas; a redução da interferência humana; a obtenção de resultados mais confiáveis; e a geração de bancos de dados e imagens para armazenamento e consulta futura (Gomes-Junior, 2020).

O GroundEye é um sistema de análise computadorizada, que possibilita a obtenção de informações sobre uma série de variáveis, dentre elas, aquelas relacionadas ao desempenho de plântulas como comprimentos da parte aérea, sistema radicular e total da plântula e os índices de crescimento, uniformidade e vigor (Brandani et al, 2023; Manual Groundeye®, 2016). As vantagens desse sistema é a rapidez na execução da análise para grandes números de lotes e obtenção dos resultados, armazenamento de imagens para elaboração de relatórios e observações futuras (Manual Groundeye®, 2016). Torna-se interessante, portanto, avaliar a eficiência para sementes de algodão, pois já existem estudos para sementes de soja (Acha e Vieira, 2020; Brandani et al, 2023), milho (Catão et al., 2020), café (Abreu et al., 2016) e de hortaliças como repolho e salsa (Vasconcelos et al., 2018).

Outro sistema de análise computadorizada promissor é o Ilastik, que é um *software* de acesso livre, baseado em aprendizado interativo de máquina para definir modelos que permitem aos usuários analisarem e classificar imagens (Berg, 2019). Este *software* pode ser empregado em programas de controle de qualidade, para auxiliar no ranqueamento de lotes em diferentes níveis de potencial fisiológico por meio da classificação das imagens de sementes e plântulas (Limão et al., 2023; Medeiros et al., 2020a). Limão et al. (2023) verificaram que os dados obtidos pelo Ilastik permitam classificação do vigor de sementes de lentilha aos quatro dias de germinação. . Outros estudos similares comprovaram a eficiência deste *software* em sementes de soja (Medeiros et al., 2020a), arroz (Yasam et al., 2022), braquiária (Souza et al., 2023) e pinhão-manso (Medeiros et al., 2020b).

Em síntese, esses avanços tecnológicos baseados na análise de imagens de sementes e plântulas têm impulsionado o desenvolvimento de métodos mais rápidos, precisos e, em muitos casos, não destrutivos, como o teste de raios X, que permite visualizar a estrutura interna das sementes. Da mesma forma, sistemas automatizados como o GroundEye e o Ilastik possibilitam obter informações detalhadas sobre o desempenho e as características morfológicas de plântulas por meio da captura e processamento de imagens. Em conjunto, essas tecnologias, aliadas a modelos modernos de machine learning, tornam possível avaliar atributos físicos e fisiológicos

de lotes de sementes em tempo real, de forma objetiva e altamente reprodutível, ampliando significativamente o potencial das análises tradicionais (Gomes-Junior, 2020).

Neste contexto, o objetivo deste estudo foi avaliar a eficiência de imagens radiográficas, análise computadorizada de plântulas e aprendizado interativo de máquina baseado no desempenho de plântulas para a classificação do potencial fisiológico de sementes de algodão.

MATERIAL E MÉTODOS

A pesquisa foi conduzida no Laboratório de Pesquisa de Sementes do Departamento de Agronomia da Universidade Federal de Viçosa, em Viçosa, MG, Brasil. Foram utilizados oito lotes de algodão da cultivar 21065TLP.

Inicialmente, as sementes de cada lote foram submetidas aos seguintes testes para a caracterização da sua qualidade inicial:

Grau de umidade: Determinado em estufa a 105°C, por 24 h, segundo as Regras para Análise de Sementes (Brasil, 2025). Foram utilizadas duas repetições com 25 sementes para cada lote e os resultados foram expressos em porcentagem (base úmida).

Germinação: quatro repetições de 50 sementes foram distribuídas em rolos de papel toalha umedecido com água na quantidade equivalente a 2,5 vezes o peso do papel seco e mantidas em germinador a 25 °C. As avaliações foram realizadas no quarto e décimo segundo dia após a semeadura e os resultados expressos em porcentagem de plântulas normais obtidas no décimo dia (Brasil, 2025).

Primeira contagem de germinação: foi realizado juntamente com o teste de germinação, consistindo na porcentagem de plântulas normais obtidas no quarto dia após a semeadura (Krzyzanowski et al., 2020).

Germinação a baixa temperatura: A metodologia utilizada foi semelhante à descrita para o teste de germinação, porém os rolos foram colocados em sacos plásticos e mantidos em câmara incubadora (tipo BOD) à temperatura de 18 °C, sendo realizada uma única avaliação no sétimo dia após a semeadura (Dias e Marcos-Filho, 2020). O resultado foi expresso em porcentagem de plântulas normais com tamanho maior ou igual a 4,0 cm.

Envelhecimento acelerado: 250 sementes de cada lote foram distribuídas sobre tela metálica acoplada em caixa plástica tipo gerbox contendo ao fundo 40 mL de água. As caixas foram tampadas de modo a se obter aproximadamente 100% UR em seu interior e mantidas em BOD a 41 °C por 72 h. Decorrido esse período, foi conduzido o teste de germinação conforme descrito acima, obtendo-se a porcentagem de plântulas normais no quarto dia após a semeadura (Marcos-Filho, 2020b).

Emergência de plântulas e Índice de velocidade de emergência (IVE): foi realizado em sala de crescimento, utilizando-se bandejas plásticas contendo mistura de solo e areia na proporção de 1:1, com quatro repetições de 50 sementes semeadas na profundidade de 1,0 cm. Foram realizadas contagens diárias e a porcentagem de emergência foi calculada computando-se o total de plântulas emergidas, com as plúmulas acima da superfície do substrato, até a completa estabilização do estande (Krzyzanowski et al., 2020). O *IVE* foi calculado conforme Maguire (1962).

Análises radiográficas

Raios X – foram utilizadas oito repetições de 25 sementes de cada lote. As sementes foram posicionadas com o eixo embrionário voltado para baixo e fixadas de forma ordenada sobre papel adesivo, para possibilitar posterior identificação individual nas análises subsequentes. Em seguida, as sementes foram colocadas no interior do equipamento de raios X digital, modelo MX-20 (Faxitron X-ray Corp. Wheeling, IL, U.S.A). Para gerar as imagens radiográficas, o equipamento foi configurado com o tempo de exposição à radiação de 15 segundos, tensão de 29 kV, distância focal de 34,6 cm e contraste da imagem calibrado em 6884 (*width*) x 3847 (*center*). As imagens radiográficas foram processadas e avaliadas utilizando-se o software ImageJ, por meio da ferramenta IJCropSeed (Medeiros et al., 2020). Após o processamento, foram extraídas as informações relacionadas às seguintes variáveis:

Área - Seleção obtida em pixels quadrados e posteriormente em unidades de milímetros quadrados (mm²).

Perímetro - O comprimento do limite externo da seleção, em milímetros (mm).

Preenchimento - Porcentagem da semente totalmente preenchida por material de alta densidade (%).

Densidade Relativa - Média dos valores cinza de todos os pixels na área de seleção (cinza.mm⁻¹).

Densidade Integrada - Produto de área e pela escala de cinza média (cinza.mm².pixel⁻¹).

Escala de cinza – É a mediana da escala de cinza de todos os pixels na área de seleção (cinza.pixel⁻¹).

Análise de imagem baseada no desempenho de plântulas pelo GroundEye

Inicialmente foi montado o teste de crescimento de plântulas, com 8 repetições de 20 sementes por lote. As sementes foram distribuídas em uma linha traçada no terço superior do papel de germinação no sentido longitudinal. As sementes foram posicionadas de forma que o

hilo ficasse direcionado para a parte inferior do papel e os rolos foram acondicionados em sacos plásticos, colocados verticalmente no germinador nos períodos de 2, 3 e 4 dias a 25 °C. Ao final desses períodos, as plântulas foram transferidas do papel de germinação para a bandeja de acrílico transparente acoplada no equipamento, de forma que não se tocassem. A captura das imagens foi realizada por uma câmera profissional de alta resolução acoplada ao equipamento GroundEye.

O processamento das imagens foi realizado por meio do software GroundEye obtendo-se os valores de comprimento do hipocótilo, comprimento da raiz e comprimento total das plântulas de algodão, bem como os valores de índice de crescimento, uniformidade e vigor conforme a equação abaixo:

$$\text{Índice de Crescimento} = \text{mín} (w_h * l_h + w_r * l_r, 1000),$$

$$\text{Índice de Uniformidade} = \text{máx} (1000 - (w_{Sh} * s_h + w_{Sr} * s_r + w_{Stotal} * s_{total} + w_{Sr/h} * s_{r/h}) - W_m * n^{\circ} \text{ sementes mortas}, 0),$$

$$\text{Índice de Vigor} = pC \text{ crescimento} + pU \text{ uniformidade},$$

Sendo,

- l_h e l_r : médias amostrais do comprimento do hipocótilo e da raiz;
- s_h, s_r, s_{total} e $s_{r/h}$: desvio padrão do comprimento do hipocótilo, da raiz, total e da razão da raiz pelo hipocótilo;
- pC e pU : peso dos índices de crescimento e uniformidade;
- w : o peso dos parâmetros

Classificação de plântulas por meio do software Ilastik

As imagens obtidas conforme descrito acima para o GroundEye, foram carregadas no software *Ilastik*. Para a segmentação, foi utilizada a ferramenta de classificação de pixels, definindo duas classes de segmentação: “semente ou plântula” (região de interesse) e “fundo” (região a ser descartada). A probabilidade de o pixel pertencer às classes de segmentação semântica foi estimada para cada pixel da imagem e ilustradas na Figura 2. Em seguida, foi criado um modelo para o reconhecimento de quatro classes de plântulas: (a) plântulas normais fortes (aquelas morfologicamente bem formadas, sem defeitos, com crescimento vigoroso da parte aérea e raiz), (b) plântulas normais fracas (aquelas morfologicamente bem formadas, mas com pequenos defeitos que ainda as caracterizam como normais ou com menor tamanho que as normais fortes), (c) plântulas anormais (com defeitos graves, como ausência ou deformação de estruturas essenciais, conforme Brasil (2025) há deformação ou ausência de alguma estrutura essencial) e (d) sementes mortas, conforme o padrão ilustrado na Figura 1 (Limão et al., 2023;

Souza et al., 2023). O classificador treinado foi aplicado para todas as imagens, gerando dados de número de plântulas normais fortes, plântulas normais fracas, plântulas anormais e sementes mortas. Um breve resumo das principais etapas do método interativo para classificação de sementes e plântulas de algodão é ilustrado na Figura 2.

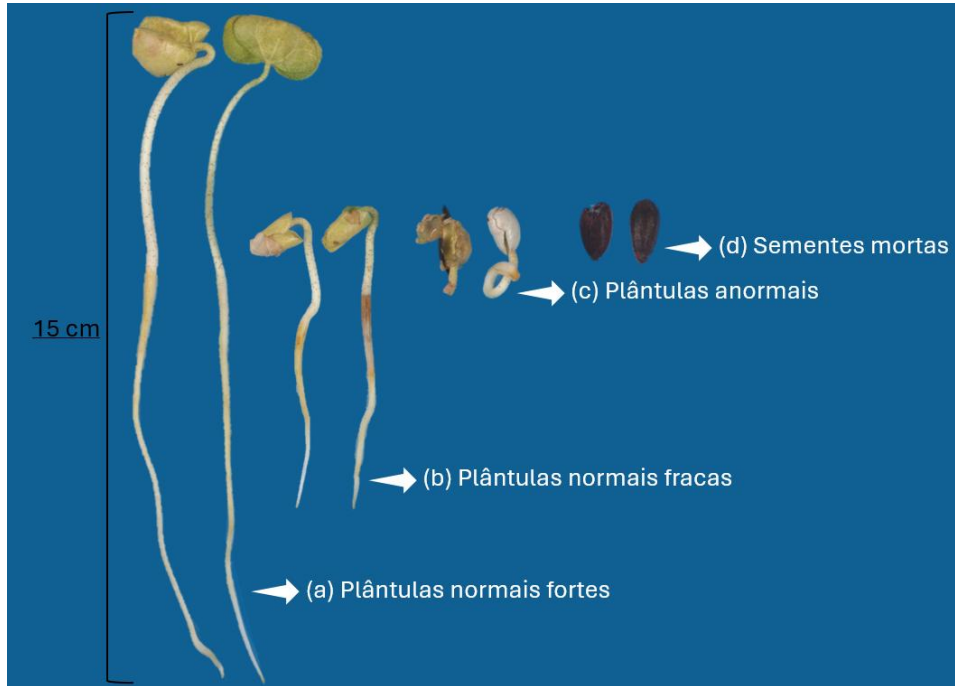


Figura 1. Padrão de plântulas normais fortes, plântulas normais fracas, plântulas anormais e sementes mortas de algodão pelo modelo de classificação utilizando o software Ilastik.

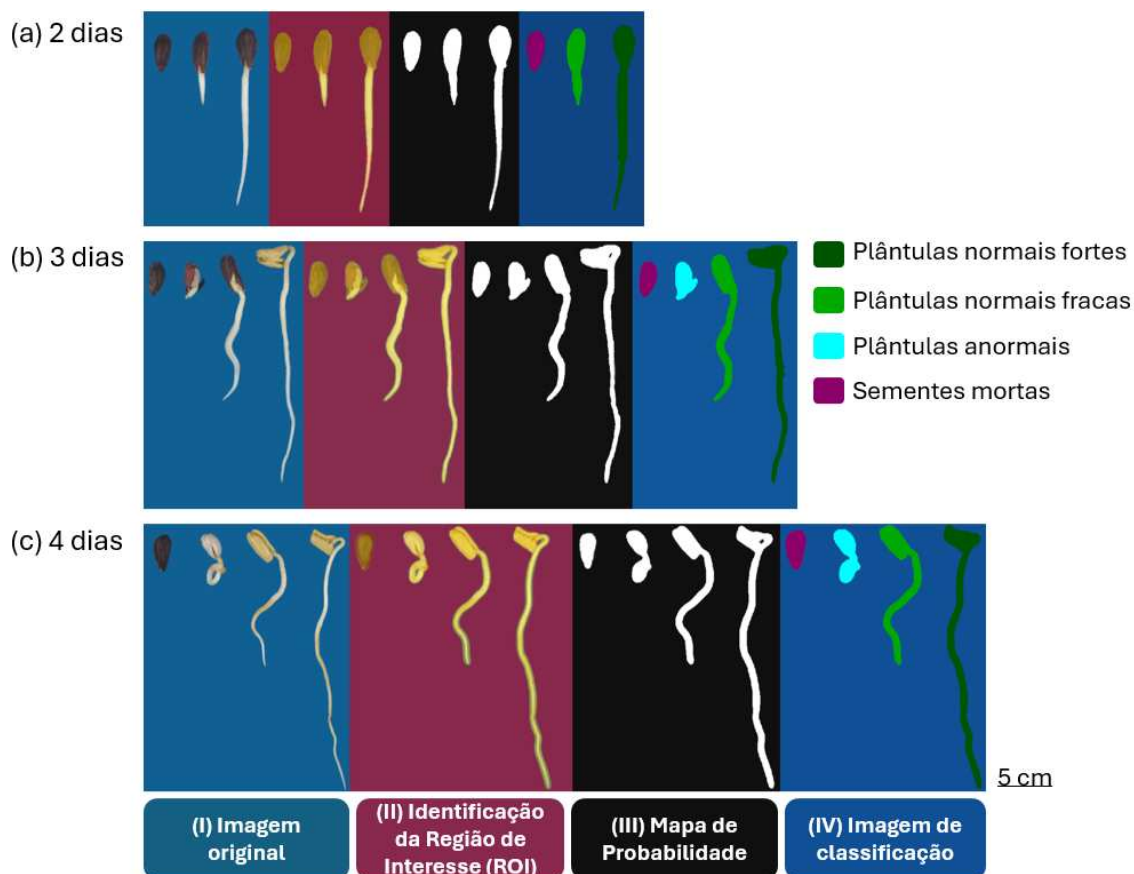


Figura 2. Esquema representativo das etapas de treinamento e classificação do aprendizado de máquina interativo para o potencial fisiológico de diferentes lotes de algodão aos 2 (a), 3 (b) e 4 (c) dias após a semeadura. Plântulas normais fortes (cor verde escuro), plântulas normais fracas (cor verde claro), plântulas anormais (cor azul neon) e sementes mortas (cor roxo). Imagens originais das plântulas sobre o fundo azul (I). Em seguida, a renderização foi aplicada para segmentação e identificação da região de interesse (ROI) para obtenção do mapa de probabilidade (II). Após a identificação da ROI, o software identifica plântulas e sementes individuais usando o mapa de probabilidade (III) e, finalmente, a predição da classificação para cada plântula ou semente é feita com base nas cores estabelecidas para treinamento (IV).

Delineamento experimental e análise estatística

O ensaio foi conduzido em delineamento inteiramente casualizado (DIC), com quatro repetições. Os dados foram testados quanto à distribuição normal dos erros pelo teste de Shapiro-Wilk e à homogeneidade de variâncias pelo teste de Bartlett. Em seguida, os dados foram submetidos à análise de variância e as médias obtidas para cada lote foram comparadas pelo teste de Tukey ($p < 0,05$).

Também foi realizada a análise de componentes principais (PCA) utilizando-se os dados obtidos para todos os testes de caracterização do potencial fisiológico inicial, para as variáveis

das imagens radiográficas e do teste de crescimento e classificação de plântulas. Para todas as análises, foi utilizado o software estatístico R 4.3.3. (R Core Team, 2025).

RESULTADOS E DISCUSSÃO

Caracterização da qualidade fisiológica

O grau de umidade dos lotes de sementes variou de 9,4 a 10,3% (Tabela 1). Trata-se de uma característica fundamental a ser considerada durante a condução dos testes, assegurando a padronização das avaliações e a obtenção de resultados consistentes (Marcos-Filho, 2015).

Tabela 1. Grau de umidade (U); germinação (G); primeira contagem de germinação (PCG); germinação em baixa temperatura (GBT); envelhecimento acelerado (EA); emergência (E) e índice de velocidade de emergência (IVE) de oito lotes de sementes da cultivar de algodão 21065TLP.

| Lotes | U (%) | G (%) | PCG (%) | GBT (%) | EA (%) | E (%) | IVE (índice) |
|--------|----------|----------|------------|------------|-----------|----------|-----------------|
| 1 | 10,1 | 96 a | 88 a | 86 a | 83 a | 93 a | 8.4 a |
| 2 | 9,8 | 90 b | 83 b | 83 ab | 81 ab | 90 ab | 7.7 b |
| 3 | 9,4 | 89 b | 80 b | 80 b | 78 b | 89 ab | 7.4 bc |
| 4 | 9,4 | 82 c | 77 c | 72 c | 69 c | 86 c | 6.5 c |
| 5 | 10,3 | 95 a | 90 a | 87 a | 84 a | 91 a | 8.5 a |
| 6 | 9,5 | 89 b | 82 b | 83 ab | 79 b | 89 ab | 7.9 b |
| 7 | 9,7 | 87 b | 80 b | 82 ab | 69 c | 87 b | 7.6 b |
| 8 | 9,7 | 79 c | 78 c | 80 b | 35 d | 84 c | 7.3 bc |
| CV (%) | - | 1,8 | 2,47 | 3,18 | 3,23 | 2,1 | 2,31 |

CV = coeficiente de variação. Médias seguidas pela mesma letra em uma coluna não diferem entre si de acordo com o teste de Tukey ($p < 0,05$).

A avaliação do potencial fisiológico dos lotes evidenciou que os lotes 1 e 5 apresentaram o melhor desempenho nos testes de germinação, primeira contagem de germinação, germinação a baixa temperatura, envelhecimento acelerado, emergência e índice de velocidade de emergência (Tabela 1). Em contraste, os lotes 4 e 8 apresentaram os piores desempenhos. No teste de germinação em baixa temperatura (GBT), os lotes 1 e 5 mantiveram o melhor desempenho, sem diferir significativamente dos lotes 2, 6 e 7, mas com valores superiores aos observados nos lotes 4 e 8, destacando-se o lote 4 como o de pior desempenho.

Como observado no presente estudo, os lotes apresentaram diferenças quanto ao potencial fisiológico. A avaliação do potencial fisiológico das sementes é fundamental para estimar o desempenho dos lotes no campo e no armazenamento; portanto, o uso em conjunto de mais de um teste permite identificar diferenças quanto ao vigor das sementes (Marcos-Filho, 2020a), possibilitando o agrupamento de lotes em classes com base em seu desempenho (Tabela 1), o que é fundamental no controle de qualidade nas empresas de sementes.

Análises radiográficas

Pelos resultados das variáveis obtidas por meio das imagens radiográficas como área (Figura 3a), perímetro (Figura 3b) e preenchimento (Figura 3c), que estão diretamente relacionadas ao tamanho das sementes e à integridade do tegumento e do embrião, a ausência de diferença estatística entre os lotes indica que os lotes apresentam elevada uniformidade física. Esse resultado sugere que todos os lotes possuem alta qualidade física, com sementes bem formadas e sem danos estruturais aparentes, o que dificulta associar estas variáveis à possíveis diferenças no potencial fisiológico das sementes (Tabela 1). Nesse contexto, a técnica de raios X se destaca como uma ferramenta altamente eficiente e não destrutiva para avaliação de características físicas das sementes, relacionadas ao tegumento e também aos tecidos internos, permitindo identificar má formação, danos mecânicos ou tecidos deteriorados. No entanto, quando os lotes apresentam padrão físico semelhante e elevado, como observado neste estudo, a técnica pode não evidenciar diferenças, não por limitação metodológica, mas sim pela homogeneidade e boa qualidade dos materiais avaliados.

Além disso, reforça-se que características físicas, especialmente tamanho da semente, nem sempre estão diretamente associadas ao desempenho fisiológico, como germinação e vigor, conforme já demonstrado por Marcos-Filho et al. (2000). Dessa forma, embora a análise radiográfica seja importante para avaliar a integridade estrutural das sementes, sua associação com resultados de testes fisiológicos permite uma avaliação mais completa da qualidade das sementes.

Para as variáveis de densidade relativa (Figura 3d), densidade integrada (Figura 3e) e média de cinzas (Figura 3f) relacionadas às estruturas internas das sementes, foi possível distinguir os lotes que apresentaram menor potencial fisiológico daqueles de alto e médio. Estudos com crambe (Ribeiro et al., 2021), brócolis (Abud et al., 2018) e leucena (Medeiros et al., 2018) mostraram que o uso dos parâmetros de densidade tecidual e média de cinzas obtidas a partir das imagens radiográficas foi promissor e contém informações relevantes acerca da

qualidade fisiológica das sementes, uma vez que sementes com menor densidade apresentaram menor vigor, originando plântulas anormais, conforme observado na Figura 4.

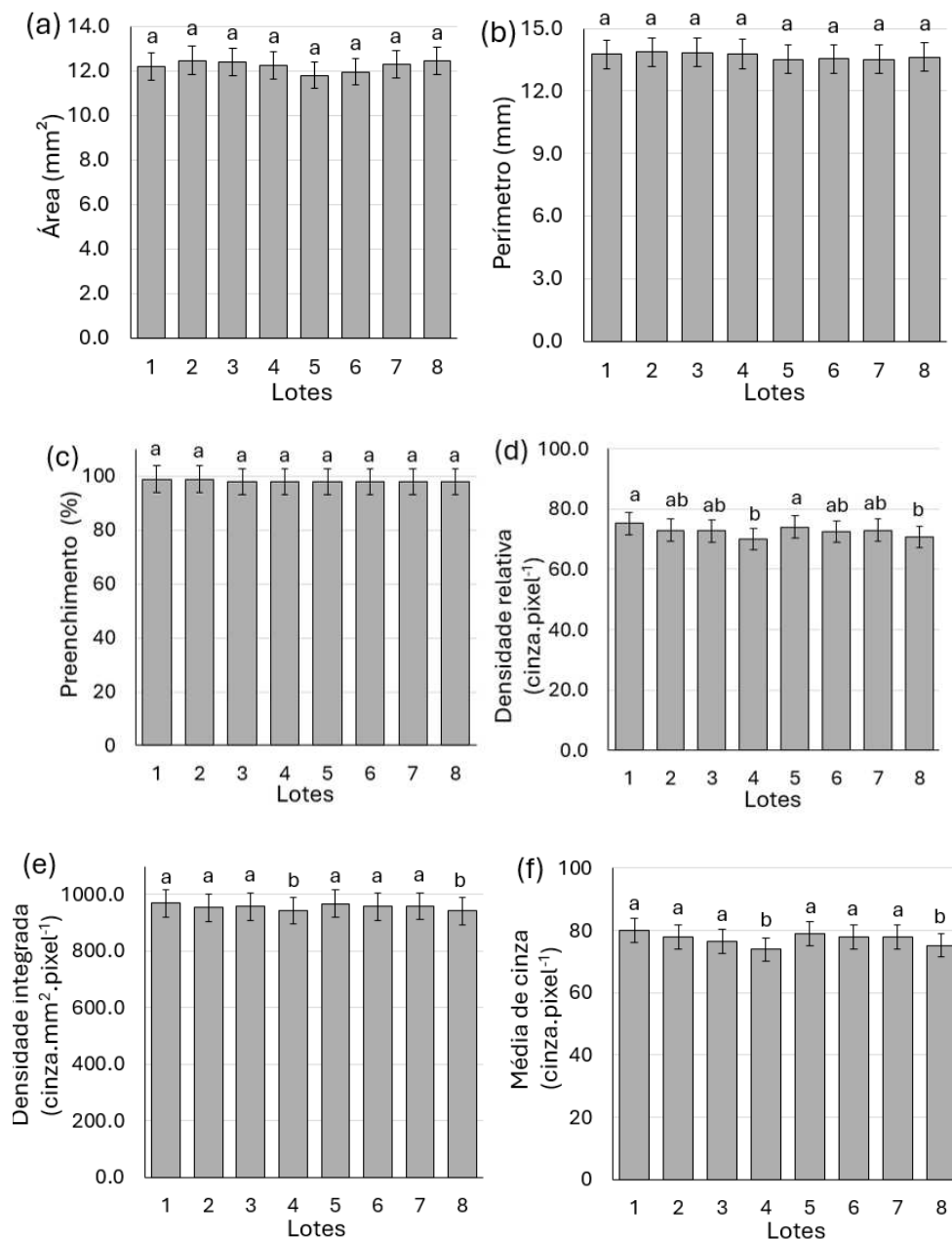


Figura 3. Valores de área (a), perímetro (b), preenchimento (c), densidade relativa (d), densidade integrada (e) e média de cinza (f) para oito lotes de sementes de algodão. Letras iguais não diferem pelo teste de Tukey ($p < 0,05$). Barras representam o desvio padrão.

Desse modo, é possível ressaltar que as variáveis de densidade relativa, densidade integrada e média de cinza pode estar relacionadas com o potencial fisiológico das sementes de algodão, conforme observado na Figura 5. Esta técnica já foi utilizada para avaliar atributos

morfométricos e integridade do embrião associados à qualidade fisiológica de sementes de diversas espécies, como soja (Wendt et al., 2014), crambe (Ribeiro et al., 2021), quiabo (Santos et al., 2020), brócolis (Abud et al., 2018), mamão (Nogueira et al., 2024) e leucena (Medeiros et al., 2018). O uso do teste de raios X na análise de sementes de algodão possibilitou identificar sementes intactas e sem nenhum tipo de dano no tegumento ou embrião (Figura 4a) e separá-las daquelas com algum tipo de injúria ou danos no tegumento, o que culmina em danos no embrião (Figura 4b).

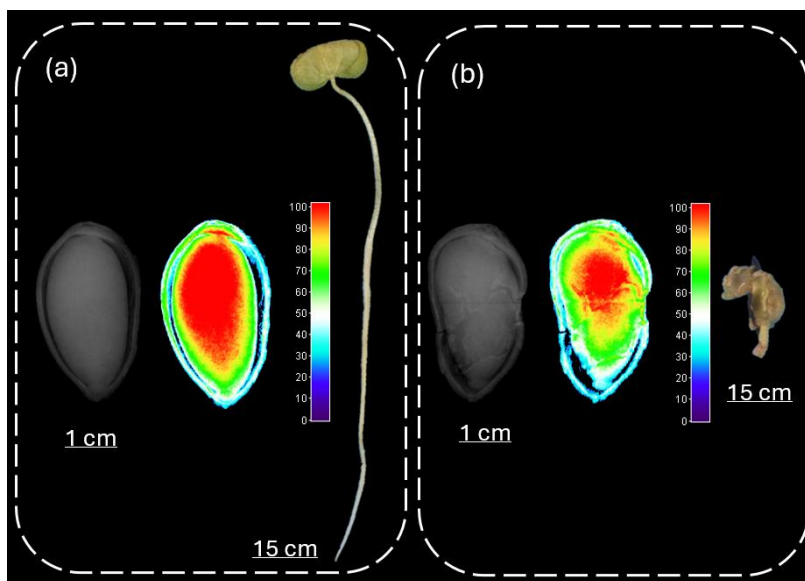


Figura 4. Imagem radiográfica da semente, mapa de calor baseado na densidade relativa da semente e imagem da plântula obtida de cada semente de algodão (a – sem danos e com embrião bem definido e formado; b – trincas no tegumento e danos no embrião).

Em termos técnicos, os raios X atravessam a semente e sofrem atenuação proporcional à densidade dos tecidos: regiões mais densas (como embriões bem formados e tecidos de reserva íntegros) absorvem mais radiação e aparecem mais claras na imagem, enquanto regiões menos densas (como espaços vazios, danos mecânicos ou tecidos deteriorados) aparecem mais escuras (Figura 4a e 4b). Assim, a densidade radiográfica funciona como um indicador indireto da organização interna da semente. Sementes com maior densidade apresentam melhor formação do embrião e maior quantidade de reservas, o que favorece o metabolismo inicial da germinação, garantindo energia e substratos para o crescimento da plântula; por outro lado, sementes com menor densidade frequentemente apresentam deterioração, desorganização celular, danos no embrião ou presença de espaços vazios, fatores que comprometem processos fisiológicos essenciais (Campos et al., 2022). Além disso, a deterioração está associada a alterações bioquímicas, como degradação de membranas, perda de compartimentalização

celular e consumo ou degradação das reservas. Esses processos reduzem a densidade interna e, conseqüentemente, a capacidade de absorção de raios X. Como resultado, sementes menos densas tendem a apresentar menor germinação e vigor, estabelecendo a relação entre densidade radiográfica e qualidade fisiológica (Campos et al., 2022).

Conforme observado na análise de componentes principais (Figura 5), é possível constatar relação direta dos vetores associados às variáveis fisiológicas, densidade relativa, densidade integrada e média de cinzas com os lotes 1 e 5, que apresentaram o maior potencial fisiológico, conforme observado na Tabela 1 e Figura 3d, 3e e 3f. Em contrapartida, os lotes 4 e 8 encontram-se nos *scores* negativos, não estando relacionados às respectivas variáveis fisiológicas e físicas.

De forma geral, a combinação dessas variáveis morfométricas permite identificar sementes bem formadas e a integridade dos tecidos. O uso da análise de raios X constitui-se como um teste complementar aos métodos tradicionais de avaliação da qualidade fisiológica, fornecendo informações sobre a morfologia interna e o estado de integridade das sementes de algodão (Cícero; Van Der Heijden, 2020; Medeiros et al., 2023). Esse tipo de avaliação é altamente vantajoso para empresas produtoras de sementes no controle de qualidade de lotes, visto que, por inspeção visual, não é possível identificar tais danos e diferenciá-los, além da redução do tempo de análise e da eficiência na obtenção de informações.

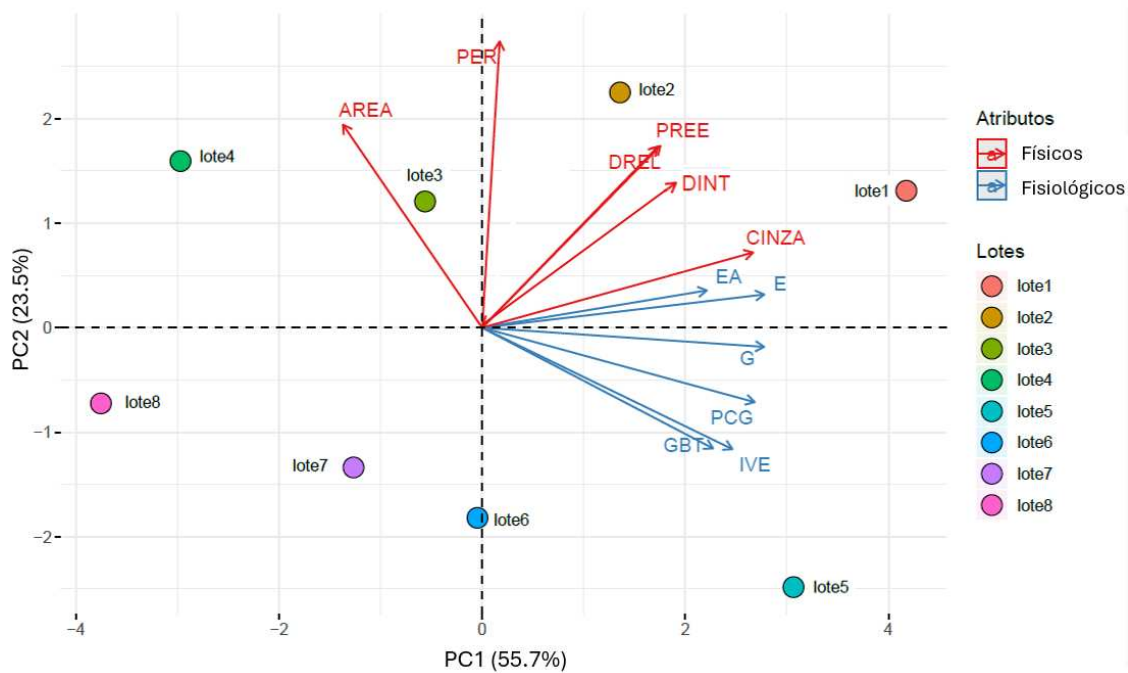


Figura 5. Análise de componentes principais das variáveis fisiológicas: germinação (G), primeira contagem de germinação (PCG), germinação a baixa temperatura (GBT),

envelhecimento acelerado (EA), emergência (E) e índice de velocidade de emergência (IVE), e das variáveis obtidas no teste de raios X: área (AREA), perímetro (PER), preenchimento (PREE), densidade relativa (DREL), densidade integrada (DINT) e média de cinza (CINZA) dos oito lotes de sementes de algodão.

Análise de imagem baseada no desempenho de plântulas por meio do software GroundEye

Os resultados da análise computadorizada de plântulas por meio do *software* GroundEye possibilitaram observar diferenças significativas no comprimento da parte aérea (CPA) das plântulas dos diferentes lotes a partir do segundo dia de germinação, sendo que os lotes 1 e 5 apresentaram melhor desempenho, os lotes 2, 3, 4 e 8, o menor crescimento da parte aérea, enquanto os lotes 6 e 7 ficaram com valores intermediários (Tabela 2). No entanto, observa-se que no terceiro dia de germinação foi possível separar os lotes em três níveis quanto ao CPA, com destaque para os lotes 1, 5 e 6 como os de maiores valores de CPA, e menores valores para os lotes 4 e 8, conforme já observado na avaliação feita no segundo dia (Tabela 2). É importante observar que, no quarto dia, foi possível classificar os lotes em apenas dois níveis de vigor, o que indica o tempo maior para as plântulas se desenvolverem, destacando-se os lotes 1 e 5 como os de maior CPA. Curiosamente, os lotes 2, 3, 6 e 7 não diferiram estatisticamente dos lotes 1 e 5, e estes, por sua vez, também não diferiram estatisticamente dos lotes 4 e 8 (Tabela 2).

Para o comprimento radicular (CR), observa-se que os lotes 1, 5 e 6 foram superiores aos demais, e os lotes 4 e 8 apresentaram os menores resultados, ficando os lotes 2, 3 e 7 em posição intermediária aos dois dias de germinação (Tabela 2). No terceiro dia, os lotes 1, 5 e 6 mantiveram-se no *ranking* como os de maior CR, e apenas o lote 8 como o de menor CR, ficando os demais lotes em posição intermediária. No quarto dia, observa-se que apenas os lotes 1 e 5 tiveram maior CR, e os lotes 4 e 8 menores CR, conforme observado aos dois dias de germinação, com os demais lotes em posições intermediárias (Tabela 2).

Em relação ao comprimento total de plântulas (CTP), observa-se que, aos dois dias, foi possível separar os lotes em três níveis de potencial fisiológico, com destaque para o maior desempenho dos lotes 1 e 5, e o menor desempenho para os lotes 4 e 8, ficando os lotes 2, 3, 6 e 7 em posições intermediárias (Tabela 2). Aos três dias, observa-se uma separação ainda maior entre os níveis de vigor; entretanto, o comportamento dos lotes permanece o mesmo, com destaque de maior CTP para os lotes 1 e 5, e menor CTP para os lotes 4 e 8, ficando os demais lotes em posições intermediárias (Tabela 2). Aos quatro dias, os lotes 1 e 5 permanecem com o

maior desempenho, já o lote 8 se destaca como o de menor CTP, seguido dos lotes 4 e 7 (Tabela 2). Considerando que a rapidez na obtenção dos resultados dos testes de vigor é um fator importante na tomada de decisão quanto ao manejo e destino dos lotes, recomenda-se a realização das avaliações de CPA, CR e CTP aos 3 dias após a semeadura, uma vez que permitiram a classificação dos lotes em diferentes níveis de potencial fisiológico.

Sementes mais vigorosas possuem maior acúmulo de reservas, atividade metabólica e integridade estrutural, o que permite germinação e crescimento inicial das plântulas de forma rápida. Logo, as características associadas ao crescimento de plântulas são frequentemente utilizadas como indicadores de potencial fisiológico, refletindo a verdadeira capacidade do estabelecimento das plântulas no campo de forma rápida e uniforme (Krzyzanowski et al., 2020). O Groundeye mostrou-se eficiente na determinação do vigor de sementes de milho (Andriuzzi et al., 2020). Em sementes de soja, os dados obtidos pelo equipamento Groundeye foram eficientes para a determinação do vigor aos dois dias, reduzindo em 6 dias o tempo necessário para a obtenção dos resultados em relação ao teste de germinação (Brandani et al., 2022). Estudos para verificar a eficiência de técnicas de análise de imagem para avaliação do potencial fisiológico também foram realizados em sementes de lentilha (Limão et al., 2023) e grão-de-bico (Araújo et al., 2021) por meio do *software* SAPL®, no qual, estes autores concluíram que a técnica é altamente precisa para avaliação do potencial fisiológico.

Os resultados do índice de uniformidade (IU), índice de crescimento (IC) e índice de vigor (IV) permitiram a classificação dos lotes em níveis vigor., destacando os melhores índices para os lotes 1 e 5 aos dois, três e quatro dias após a germinação (Tabela 2). Vale destacar que o lote 6 apresentou desempenho similar aos lotes 1 e 5 aos dois e três dias para o IU, IC e IV (Tabela 2). Os lotes 4 e 8 apresentaram os menores desempenhos no IU, IC e IV aos dois, três e quatro dias após a germinação, com a classificação similar à observada para o crescimento da parte aérea, principalmente aos dois e três dias (Tabela 2).

O IU, IC e IV se mostraram eficientes na diferenciação dos níveis de potencial fisiológico de sementes de soja (Brandani et al., 2022) e milho (Andriuzzi et al., 2020). Em sementes de lentilha (Limão et al., 2023) e grão-de-bico (Araújo et al., 2021), os dados obtidos pelo *software* SAPL foram eficientes na separação dos lotes quanto ao vigor aos 3 e 4 dias, respectivamente.

Tabela 2: Resultados médios de comprimento de parte aérea (CPA), comprimento radicular (CR), comprimento total de plântulas (CTP), índice de uniformidade (IU), índice de

crescimento (IC) e índice de vigor corrigido (IVC), obtidos por meio do GroundEye aos dois, três e quatro dias após a semeadura de oito lotes de sementes de algodão

| Lote | CPA | CR | CT | IU | IC | IV |
|---------------|------------------------------|--------|--------|----------|----------|----------|
| | (cm.plântula ⁻¹) | | | Índices | | |
| 2 dias | | | | | | |
| 1 | 1,5 a | 2,9 a | 4,4 a | 848,7 a | 278,1 a | 472,2 a |
| 2 | 1,2 c | 2,3 c | 3,5 bc | 778,3 b | 244,7 b | 417,3 b |
| 3 | 1,1 c | 2,2 cd | 3,3 bc | 758,0 b | 231,7 b | 408,1 b |
| 4 | 1,0 d | 2,1 d | 3,1 c | 725,1 c | 211,1 c | 375,3 c |
| 5 | 1,5 a | 2,9 a | 4,4 a | 856,2 a | 293,8 a | 479,8 a |
| 6 | 1,4 b | 2,9 a | 4,3 a | 843,8 a | 275,6 a | 454,2 ab |
| 7 | 1,3 bc | 2,5 b | 3,8 b | 823,5 ab | 234,6 b | 407,9 b |
| 8 | 1,1 c | 2,1 d | 3,2 c | 727,2 c | 215,8 c | 385,6 c |
| CV (%) | 2,3 | 1,6 | 1,1 | 1,2 | 3,2 | 1,3 |
| 3 dias | | | | | | |
| 1 | 3,0 a | 6,4 a | 9,4 a | 851,7 a | 595,7 a | 678,3 a |
| 2 | 2,7 b | 6,0 ab | 8,7 b | 823,2 ab | 552,9 ab | 644,5 b |
| 3 | 2,7 b | 5,9 b | 8,6 c | 815,4 ab | 544,9 ab | 636,3 b |
| 4 | 2,5 c | 5,4 b | 7,9 d | 796,14 b | 533,1 b | 630,7 c |
| 5 | 2,9 a | 6,3 a | 9,2 a | 844,9 a | 594,7 a | 680,0 a |
| 6 | 2,9 a | 6,2 a | 9,1 a | 834,4 a | 581,6 ab | 669,2 ab |
| 7 | 2,7 b | 5,5 b | 8,2 c | 828,9 ab | 559,0 ab | 633,3 b |
| 8 | 2,6 bc | 5,3 c | 7,9 d | 803,7 b | 536,9 b | 614,7 c |
| CV (%) | 1,6 | 3,0 | 1,9 | 1,3 | 2,1 | 1,8 |
| 4 dias | | | | | | |
| 1 | 5,7 a | 9,7 a | 15,4 a | 886,3 a | 892,7 a | 854,7 a |
| 2 | 5,4 ab | 8,6 b | 14,0 b | 834,1 bc | 862,3 ab | 824,8 b |
| 3 | 5,3 ab | 8,8 b | 14,1 b | 823,7 bc | 844,9 ab | 835,4 b |
| 4 | 5,1 ab | 7,9 c | 13,0 c | 775,2 c | 804,7 b | 803,9 c |
| 5 | 5,6 a | 9,6 a | 15,2 a | 865,1 a | 883,5 a | 852,3 a |
| 6 | 5,5 ab | 9,3 a | 14,8 b | 854,5 b | 865,0 ab | 844,9 a |
| 7 | 5,3 ab | 7,8 c | 13,1 c | 812,9 bc | 829,3 ab | 823,9 b |
| 8 | 4,9 b | 7,7 c | 12,6 d | 753,0 c | 801,7 b | 807,1 c |

| | | | | | | |
|--------|-----|-----|-----|-----|-----|-----|
| CV (%) | 2,5 | 1,3 | 1,8 | 1,6 | 1,4 | 1,2 |
|--------|-----|-----|-----|-----|-----|-----|

* = significativo pelo teste F a 5% de probabilidade; CV = coeficiente de variação. Médias seguidas pela mesma letra, na coluna, não diferem entre si pelo teste de Tukey ($p < 0,05$).

O índice de uniformidade, de crescimento e de vigor são fundamentais para a estimativa da qualidade fisiológica de sementes. Esses índices, derivados da avaliação do teste de crescimento de plântulas, constituem-se em uma ferramenta fundamental para aplicação em programas de controle de qualidade de sementes (Pereira et al., 2020; Silva et al., 2017).

Assim, a análise computadorizada de imagens de plântulas apresenta inúmeras vantagens em relação a outros testes de vigor, sendo a principal a redução da subjetividade nos resultados, uma vez que a análise automatizada permite maior padronização, precisão e rapidez nas avaliações do comprimento de plântulas (Gomes-Junior, 2020).

A análise de componentes principais (Figura 6) realça ainda mais o que foi exposto na Tabela 2, destacando os lotes 1 e 5 nas extremidades dos *scores* negativos da PC1, relacionados às variáveis de caracterização do potencial fisiológico inicial e pelas variáveis obtidas no teste de crescimento, bem como o lote 6, que também apresentou um bom desempenho neste teste. De forma contrária, os lotes 4 e 8 ficaram situados nos *scores* positivos da PC1, ou seja, sem relação nenhuma com as variáveis de potencial fisiológico e crescimento, ficando os demais lotes em posições intermediárias (Figura 6).

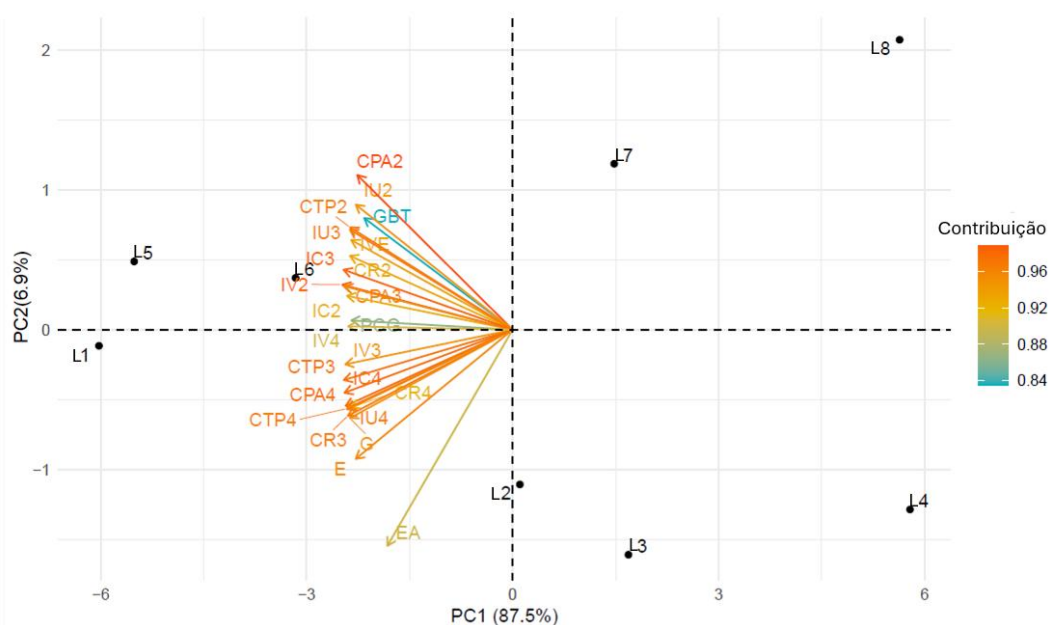


Figura 6. Análise de componentes principais das variáveis fisiológicas: germinação (G), primeira contagem de germinação (PCG), germinação a baixa temperatura (GBT),

envelhecimento acelerado (EA), emergência (E) e índice de velocidade de emergência (IVE), e das variáveis obtidas no teste de crescimento por meio do software GroundEye: comprimento da parte aérea (CPA), comprimento radicular (CR), comprimento total de plântulas (CTP), índice de uniformidade (IU), índice de crescimento (IC) e índice de vigor (IV) aos 2, 3 e 4 dias após a germinação dos oito lotes de sementes de algodão.

Classificação de plântulas por meio do software Ilastik

A porcentagem de plântulas normais fortes, plântulas normais fracas, plântulas anormais e sementes mortas geradas pelos lotes aos 2, 3 e 4 dias após a semeadura estão ilustradas na figura 7. Os lotes mais vigorosos apresentaram maior porcentagem de plântulas normais fortes, com parte aérea e raiz primária bem desenvolvidas. Em contrapartida, lotes menos vigorosos apresentaram maior porcentagem de plântulas normais fracas, plântulas anormais e sementes mortas.

No entanto, observa-se maior confiabilidade e sensibilidade na distinção entre as classes a partir do terceiro dia de germinação, resultando no ranqueamento dos lotes conforme os níveis de potencial fisiológico, com destaque para a maior porcentagem de plântulas normais fortes para os lotes 1 e 5, seguido dos lotes 2, 6 e 7 em posição intermediária, e os lotes 3, 4 e 8 como os de pior desempenho nessa classe e maior ocorrência na classe de plântulas normais fracas e anormais (Figura 7b). Para a classe de sementes mortas, apenas o lote 4 diferiu dos demais, por não apresentar esta categoria de sementes (Figura 7b).

Aos dois dias, os lotes 1 e 5 tiveram maior porcentagem de plântulas normais fortes do que os lotes 4 e 8, não diferindo dos demais. Já aos três e quatro dias, os lotes 1 e 5 foram superiores aos demais lotes, destacando-se como inferiores os lotes 4 e 8 aos três dias e apenas o lote 4 aos quatro dias (Figura 7).

O Ilastik é um *software* de processamento de imagens baseado em aprendizado de máquina interativo que permite aos usuários classificarem e analisar imagens de forma dinâmica conforme seus objetivos (Berg, 2019). A aplicação deste *software* abre um leque de oportunidades valiosas para diversos segmentos da pesquisa científica envolvendo análise de imagens, oferecendo flexibilidade e reprodutibilidade dos resultados (Berg et al., 2019; Dietz et al., 2020). O uso de análise de imagens vem crescendo gradativamente, impulsionando assim os avanços científicos e tecnológicos para avaliação da qualidade de sementes. Limão et al. (2023) avaliaram a classificação do vigor de sementes de lentilha com base em técnicas de análise de imagens de plântulas e aprendizado interativo de máquina por meio do *software*

Ilastik aos 3, 4, 5 e 10 dias após a germinação, e concluíram que os dados obtidos pelo *software* foram sensíveis no desenvolvimento de modelos com alta precisão para avaliação do vigor aos 4 dias. Em braquiária, o *software* Ilastik foi eficiente para a classificação das sementes quanto ao potencial fisiológico com apenas 7 dias de germinação (Souza et al., 2023).

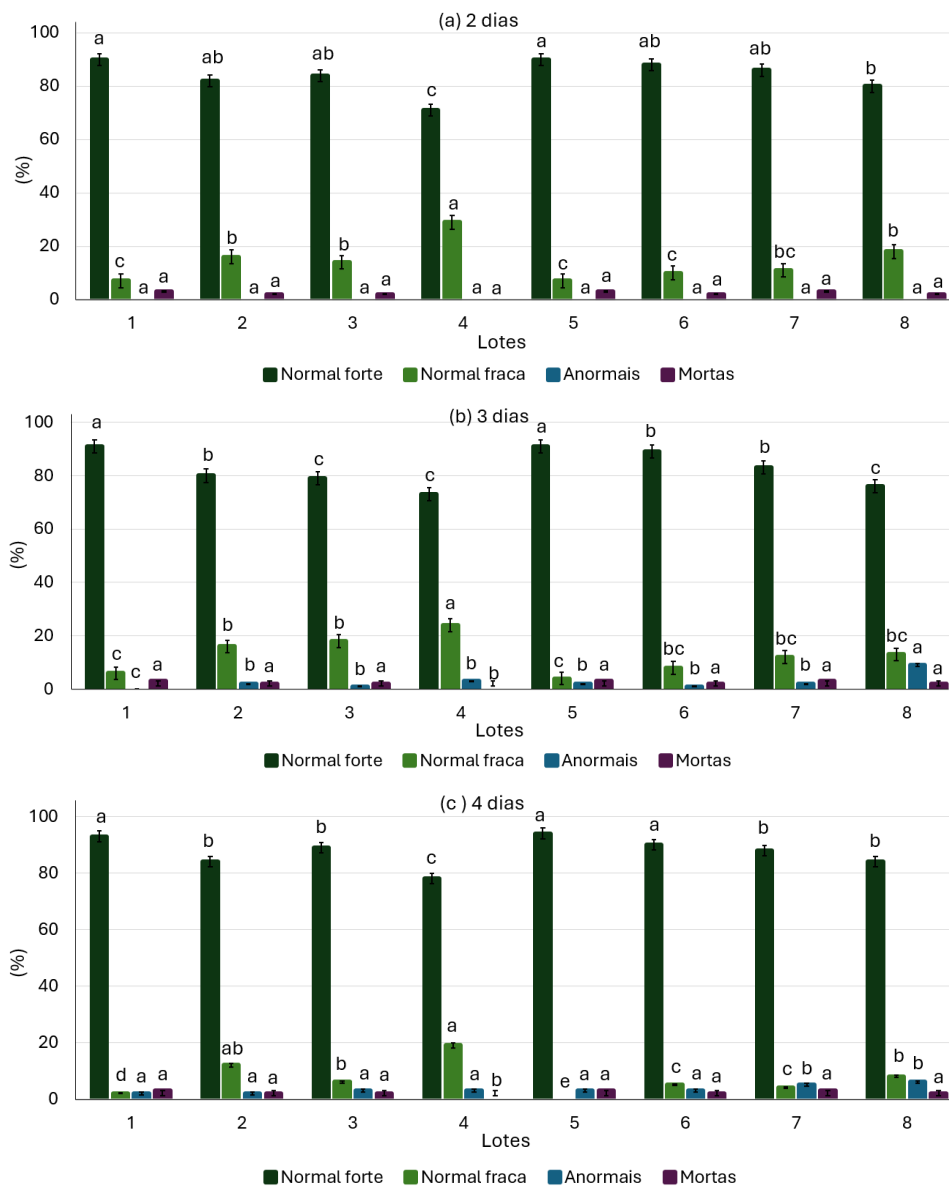


Figura 7. Porcentagem de plântulas normais fortes, normais fracas, anormais e sementes mortas aos dois (a), três (b) e quatro (c) dias após a semeadura, de acordo com os dados obtidos pelo *software* Ilastik. Letras iguais não diferem pelo teste de Tukey ($p < 0,05$). As barras representam um intervalo de confiança de 95%.

A aplicação de técnicas de aprendizado de máquina com dados gerados pelo *software* Ilastik para classificar a qualidade fisiológica de sementes e plântulas de soja foi testada por

Medeiros et al. (2020a), onde os autores concluíram que esta ferramenta foi precisa na identificação de danos nas sementes e na classificação das plântulas quanto ao vigor.

Os modelos desenvolvidos podem ser refinados e treinados no *software* Ilastik, com a possibilidade de aumentar o banco de dados (imagens) e tratamentos utilizando todas as ferramentas disponíveis e necessárias no *software*. Essas ferramentas incluem recursos de classificação rápida (*batch processing*), modelos probabilísticos e interfaces gráficas de fácil manipulação, o que possibilita um treinamento interativo e rápido, bem como a extração de informações precisas a partir de imagens (Berg et al., 2019).

A eficiência da classificação de plântulas por meio do *software* Ilastik pode ser comprovada por meio da análise de PCA (Figura 8), onde os lotes 1 e 5 estão diretamente relacionados com os vetores de potencial fisiológico nas extremidades dos *scores* negativos, seguidos dos lotes 6, 7, 2 e 3, que apresentaram desempenho intermediário. Já os lotes 4 e 8 apresentaram-se no sentido oposto, nos *scores* positivos, diretamente relacionados com os vetores de plântulas anormais e fracas aos 2, 3 e 4 dias (Figura 8).

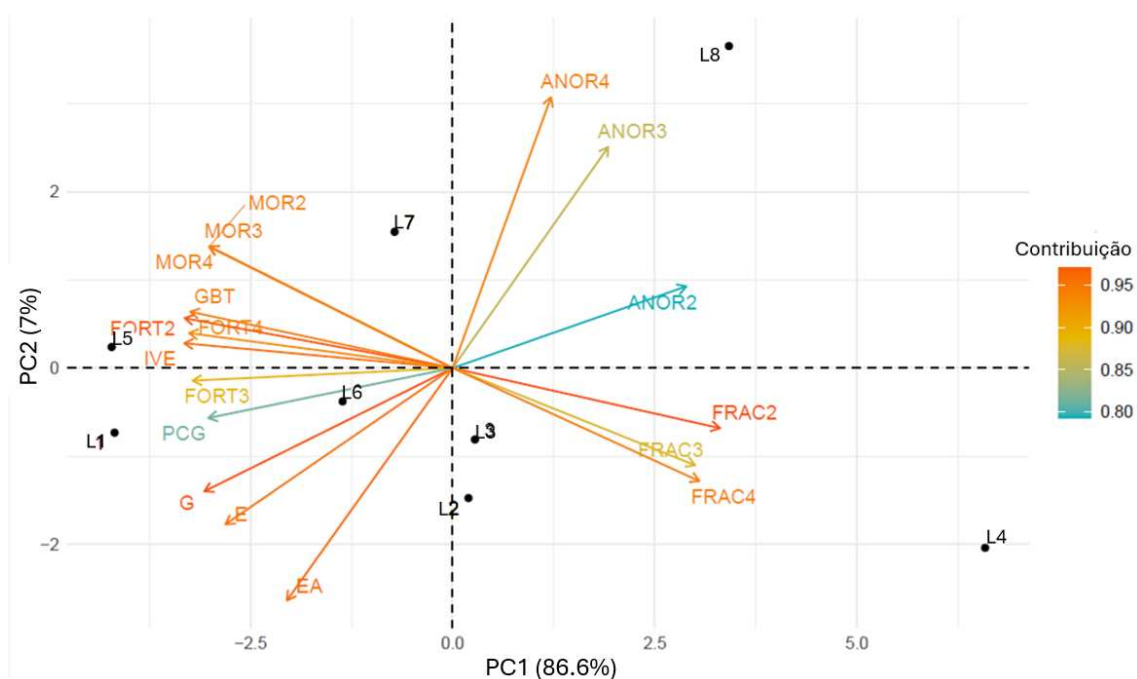


Figura 8. Análise de componentes principais das variáveis fisiológicas: germinação (G), primeira contagem de germinação (PCG), germinação a baixa temperatura (GBT), envelhecimento acelerado (EA), emergência (E) e índice de velocidade de emergência (IVE), e da classificação de plântulas obtidas pelo software Ilastik: plântulas normais fortes (FORT), plântulas normais fracas (FRAC), plântulas anormais (ANOR), sementes mortas (MOR) aos 2, 3 e 4 dias após a germinação dos oito lotes de sementes de algodão.

Os resultados do presente estudo apontam a possibilidade de aplicações práticas das técnicas de análise de imagens de sementes e de plântulas nos programas de controle de qualidade de sementes. Assim, pode-se afirmar que os parâmetros obtidos por meio da classificação das imagens geradas pelas tecnologias de imagens de raios X, análise computadorizada de plântulas pelo GroundEye e Ilastik são ferramentas promissoras para a aferição da qualidade fisiológica e vigor de sementes e são fundamentais para o desenvolvimento da agricultura moderna, uma vez que tais métodos possibilitam resultados rápidos e precisos para a classificação de sementes.

CONCLUSÕES

A técnica de raios X apresenta relação com o potencial fisiológico das sementes de algodão, além de fornecer informações sobre suas estruturas internas.

Aos três dias de germinação, o comprimento da parte aérea, da raiz primária, o comprimento total de plântula e os índices de uniformidade, crescimento e vigor, determinados por meio do GroundEye, são eficientes na classificação de lotes de algodão quanto ao potencial fisiológico.

O aprendizado de máquina supervisionado por meio do Ilastik possibilitou a classificação dos lotes quanto ao potencial fisiológico, obtendo resultados comparáveis aos testes convencionais.

REFERÊNCIAS

- ABRAPA. *Relatório de Safra Brasil*. 2025. Disponível em: <https://app.powerbi.com/view?r=eyJrIjoiNDZhMDIhZjYtZTljNi00NjNmLWIwMDEtOGY4OGE1OTE5ZmE1IiwidCI6IjRhMDk1OGIzLTg4MWQtNDBmYS05NTU1LTIwODQ1MzdhYWZkMyJ9&pageName=ReportSection10c4aa5d90c47641852d>. Acesso em: 24 set. 2025.
- ABREU, L. A. S. et al. Computerized analysis in the physiological quality of coffee seeds. *International Journal of Current Research*, v. 8, n. 11, p. 40820-40823, 2016.
- ABUD, H. F. et al. Radiographic images and relationship of the internal morphology and physiological of broccoli seed. *Acta Scientiarum: Agronomy*, v. 40, n. 1, p. 1-9, 2018.
- ACHA, A. J.; VIEIRA, H. D. Digital image processing of coated perennial-soybean seeds and correlation with physiological attributes. *Journal of Seed Science*, v. 42, e202042004, 2020.

- ANDRIAZZI, C. V. G. et al. Avaliação da qualidade fisiológica de sementes de milho por meio da análise de imagens. *Revista Agro*, v. 14, p. 1–13, 2020. DOI: 10.18227/1982-8470ragro.v14i0.6750.
- ARAÚJO, J. O. et al. Chickpea seed vigor evaluated by computerized seedling analysis. *Semina: Ciências Agrárias*, v. 42, n. 1, p. 71–80, 2021. DOI: 10.5433/1679-0359.2021v42n1p71.
- BERG, S. et al. Ilastik: interactive machine learning for (bio) image analysis. *Nature Methods*, v. 16, n. 12, p. 1226-1232, 2019. DOI: 10.1038/s41592-019-0582-9.
- BRANDANI, E. B. et al. Image analysis for the evaluation of soybean seeds vigor. *Acta Agronômica*, v. 70, n. 3, p. 311-316, 2021. DOI: 10.15446/acag.v70n3.92093.
- BRASIL. Ministério da Agricultura e Pecuária. Secretaria de Defesa Agropecuária. *Regras para Análise de Sementes – RAS*. Brasília: MAPA/SDA, 2025. https://wikisda.agricultura.gov.br/pt-br/Laborat%C3%B3rios/Metodologia/Sementes/RAS_2025/cap_15_tol_fiscalizacao
- CATÃO, H. C. R. M. et al. Antioxidant activity and physiological performance of popcorn seed after saline stress and analysis of seedling images. *Ciência e Agrotecnologia*, v. 44, p. e005020, 2020. DOI: 10.1590/1413-7054202044005020.
- CÍCERO, S. M.; VAN DER HEIJDEN, G. W. A. X-ray image analysis in seed technology: applications and recent advances. *Seed Science Research*, v. 30, n. 3, p. 158–169, 2020.
- MEDEIROS, A. D. et al. IJCropSeed: An open-access tool for high-throughput analysis of crop seed radiographs. *Computers and Electronics in Agriculture*, v. 175, 105555, 2020c. DOI: 10.1016/j.compag.2020.105555.
- DIAS, C. F. S.; MARCOS-FILHO, J. Teste de germinação a baixa temperatura. In: KRZYŻANOWSKI, F. C. et al. *Vigor de sementes: conceitos e testes*. Londrina: ABRATES, 2020. p. 317-329.
- DIETZ, C. et al. Integration of the ImageJ ecosystem in the KNIME analytics platform. *Frontiers of Computer Science*, v. 2, n. 8, 2020. DOI: 10.3389/fcomp.2020.00008.
- ELMASRY, G. et al. Utilization of computer vision and multispectral imaging techniques for classification of cowpea (*Vigna unguiculata*) seeds. *Plant Methods*, v. 15, n. 24, 2019. DOI: 10.1186/s13007-019-0411-2.

FRANÇA-SILVA, F. et al. Advances in imaging technologies for soybean seed analysis. *Journal of Seed Science*, v. 45, e202345022, 2023. DOI: 10.1590/2317-1545v45274098.

GOMES-JUNIOR, F. G. Análise computadorizada de imagens de plântulas. In: KRZYZANOWSKI, F. C. et al. *Vigor de sementes: conceitos e testes*. Londrina: ABRATES, 2020. p. 141-177.

ISTA. *Regras internacionais para testes de sementes*. Zurique: International Seed Testing Association, 2004.

KRZYZANOWSKI, F. C. et al. Testes de vigor baseados em desempenho de plântulas. In: KRZYZANOWSKI, F. C. et al. *Vigor de sementes: conceitos e testes*. Londrina: ABRATES, 2020. p. 79-127.

KOTWALIWALE, N.; SINGH, K.; KALNE, A.; JHA, S. N.; SETH, N.; KAR, A. X-ray imaging methods for internal quality evaluation of agricultural produce. *Journal of Food Science and Technology*, v. 51, n. 1, p. 1-15, 2014. DOI: <https://doi.org/10.1007/s13197-011-0485-y>

LIMÃO, M. A. R. et al. Classification of lentil seed vigor based on seedling image analysis techniques and interactive machine learning. *Journal of Seed Science*, v. 45, n. 1, e20238712, 2023. DOI: 10.1590/2317-1545v45277692.

MAGUIRE, J. D. Speed of germination: aid in selection and evaluation for seedling emergence and vigor. *Crop Science*, v. 2, n. 2, p. 176, 1962. DOI: 10.2135/cropsci1962.0011183X000200020033x.

MANUAL GROUND EYE®. Lavras: Tbit Tecnologia e Sistemas, 2016. 88 p.

MARCOS-FILHO, J. *Fisiologia de sementes de plantas cultivadas*. Piracicaba: FEALQ, 2015. 660 p.

MARCOS-FILHO, J. Teste de envelhecimento acelerado. In: KRZYZANOWSKI, F. C. et al. *Vigor de sementes: conceitos e testes*. Londrina: ABRATES, 2020b. p. 182-244.

MARCOS-FILHO, J. Testes de vigor: importância e utilização. In: KRZYZANOWSKI, F. C. et al. *Vigor de sementes: conceitos e testes*. Londrina: ABRATES, 2020a. p. 17-72.

MARCOS-FILHO, J.; NOVENBRE, A. D. L. C.; CHAMMA, H. M. C. P. Tamanho da semente e o teste de envelhecimento acelerado para soja. *Scientia Agricola*, v. 57, n. 3, p. 473-482, 2000.

MEDEIROS, A. D. et al. Classificação de qualidade de sementes de pinhão-mansão utilizando imagens radiográficas e aprendizado de máquina. *Industrial Crops and Products*, v. 146, e112162, 2020b. DOI: 10.1016/j.indcrop.2020.112162.

MEDEIROS, A. D. et al. Interactive machine learning for soybean seed and seedling quality classification. *Scientific Reports*, v. 10, n. 1, 11267, 2020a. DOI: 10.1038/s41598-020-68273-y.

MEDEIROS, A. D. et al. Parameters based on X-ray images to assess the physical and physiological quality of *Leucaena leucocephala* seeds. *Ciência e Agrotecnologia*, v. 42, n. 6, p. 643-652, 2018. DOI: 10.1590/1413-70542018426023318.

MEDEIROS, A. D. et al. X-ray image analysis for non-destructive evaluation of internal seed morphology and its relationship with physiological performance. *Seed Science and Technology*, v. 51, p. 1–14, 2023.

NOGUEIRA, G. A. et al. X-rays for determination of papaya seeds viability. *Revista Ciência Agronômica*, v. 55, e20238712, 2024. DOI: 10.5935/1806-6690.20240021.

PEREIRA, M. D. et al. Processamento digital de imagens de plântulas na avaliação do vigor de sementes de *Moringa oleifera* Lam. *Ciência Florestal*, v. 30, n. 2, p. 291, 2020. DOI: 10.5902/1980509825750.

R CORE TEAM. *R: A language and environment for statistical computing*. 2025. Disponível em: <https://www.R-project.org/>.

REHMAN, A.; KAMRAN, M.; AFZAL, I. Production and processing of quality cotton seed. In: AHMAD, S.; HASANUZZAMAN, M. (org.). *Cotton production and uses: agronomy, crop protection, and postharvest technologies*. Singapore: Springer Nature, 2020. p. 547-570. DOI: 10.1007/978-981-15-1472-2.

RIBEIRO, A. S. et al. Use of radiographic images for rapid and non-destructive assessment of crambe seed quality. *Journal of Seed Science*, v. 43, e202143001, 2021.

SANTOS, R. F. Morphological and physiological changes during maturation of okra seeds evaluated through image analysis. *Scientia Agricola*, v. 77, n. 3, e20180297, 2020.

SILVA, P. P. et al. Assessment of squash seed vigor using computerized image analysis. *Journal of Seed Science*, v. 39, n. 2, p. 159-165, 2017. DOI: 10.1590/2317-1545v39n2171177.

SIMAK, M.; GUSTAFSSON, Å. X-ray photography and sensitivity in seed tests. *Hereditas*, v. 39, n. 3-4, p. 458-468, 1953.

SOUZA, L. R. et al. Near infrared spectroscopy and seedling image analysis to evaluate the physiological potential of *Urochloa decumbens* (Stapf) R.D. Webster seeds. *Revista Brasileira de Sementes*, v. 45, n. 2, p. 277-302, 2023. DOI: 10.1590/2317-1545v4527702.

VASCONCELOS, M. C. et al. Radiography and biometric analysis of broadleaf vegetable seeds. *Amazonian Journal of Agricultural and Environmental Sciences*, v. 61, 2018.

WENDT, L. et al. Avaliação do potencial fisiológico de sementes de soja por meio de imagens. *Pesquisa Agropecuária Tropical*, v. 44, n. 3, p. 280-286, 2014.

XIA, Y. et al. Recent advances in emerging techniques for non-destructive detection of seed viability: a review. *Artificial Intelligence in Agriculture*, v. 1, p. 35–47, 2019. DOI: 10.1016/j.aiia.2019.05.001.

YASAM, S.; NAIR, S. A. N.; KUMAR, K. P. S. Previsão da capacidade de germinação de sementes baseada em aprendizagem supervisionada para agricultura de precisão. *Soft Computing*, v. 26, p. 13133–13144, 2022. DOI: 10.1007/s00500-022-06910-6.

CAPÍTULO II – Explorando ferramentas espectrais em conjunto com algoritmos de aprendizado de máquina para classificação do potencial fisiológico de sementes de algodão

RESUMO

Atualmente, técnicas não destrutivas voltadas à obtenção rápida e precisa de informações sobre o potencial fisiológico de sementes têm ganhado destaque, com ênfase na espectroscopia no infravermelho próximo (NIR) e na análise de imagens multiespectrais. Este estudo teve como objetivo avaliar a eficiência dessas técnicas aliadas a algoritmos de aprendizado de máquina para a classificação do potencial fisiológico de sementes de algodão. Foram utilizados oito lotes de sementes da cultivar 21065TLP. Para a aquisição dos espectros NIR, 200 sementes por lote foram analisadas individualmente em espectrofotômetro FT-NIR, na faixa de comprimento de onda de 1.000 a 2.500 nm. Imagens multiespectrais também foram obtidas das mesmas sementes nas bandas espectrais de 395, 460, 520, 585, 620, 740, 850 e 940 nm utilizando-se um protótipo equipado com câmera digital e LEDs. Em seguida, as mesmas sementes foram submetidas ao teste de germinação, cujo os dados foram submetidos à análise de variância e as médias obtidas foram comparadas pelo teste de Tukey ($p < 0,05$). Os dados espectrais NIR originais foram pré-processados utilizando-se os métodos: *Standard Normal Variate* (SNV); *Multiplicative Scatter Correction* (MSC); derivadas de Savitzky-Golay de 1ª e 2ª ordem. Para os dados multiespectrais, foram utilizados SNV e MSC. Os algoritmos de classificação testados foram Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA), Random Forest (RF) e Redes Neurais (NN). Para o desenvolvimento dos modelos de classificação, os lotes foram classificados em três categorias com base nos resultados do teste de germinação: alto (germinação $\geq 91\%$), médio (germinação entre 85–90%) e baixo potencial fisiológico (germinação $\leq 84\%$). A espectroscopia FT-NIR foi eficiente para a classificação das sementes de algodão de acordo com seu potencial fisiológico. Todos os modelos testados foram promissores, alcançando acurácia de até 98%. Para a análise de imagens multiespectrais, o modelo obtido com o algoritmo Random Forest foi o mais eficiente para a classificação do potencial fisiológico das sementes, atingindo acurácia de 81%. Assim, tanto a espectroscopia FT-NIR como a análise de imagens multiespectrais mostraram-se promissoras para a classificação de lotes de sementes de algodão de acordo com seu potencial fisiológico.

PALAVRAS-CHAVE: *Gossypium hirsutum* L., vigor de sementes, modelos de classificação, aprendizado de máquina, inteligência artificial.

ABSTRACT

Currently, non-destructive techniques aimed at the rapid and accurate acquisition of information on seed physiological potential have gained prominence, particularly near-infrared (NIR) spectroscopy and multispectral image analysis. This study aimed to evaluate the efficiency of these techniques combined with machine learning algorithms for classifying the physiological potential of cotton seeds. Eight seed lots of the cultivar 21065TLP were used. For NIR spectral acquisition, 200 seeds per lot were individually analyzed using an FT-NIR spectrophotometer within the wavelength range of 1,000 to 2,500 nm. Multispectral images were also obtained from the same seeds at spectral bands of 395, 460, 520, 585, 620, 740, 850, and 940 nm using a prototype equipped with a digital camera and LEDs. The same seeds were subsequently subjected to the germination test, and the data were analyzed by analysis of variance, with means compared using Tukey's test ($p < 0.05$). The original NIR spectral data were preprocessed using the following methods: Standard Normal Variate (SNV), Multiplicative Scatter Correction (MSC), and first- and second-order Savitzky–Golay derivatives. For multispectral data, SNV and MSC were applied. The classification algorithms tested were Partial Least Squares Discriminant Analysis (PLS-DA), Random Forest (RF), and Neural Networks (NN). For the development of classification models, the seed lots were categorized into three classes based on germination test results: high (germination $\geq 91\%$), medium (germination between 85–90%), and low physiological potential (germination $\leq 84\%$). FT-NIR spectroscopy was efficient for classifying cotton seeds according to their physiological potential. All tested models were promising, achieving accuracy of up to 98%. For multispectral image analysis, the model developed using the Random Forest algorithm showed the best performance, reaching an accuracy of 81%. Thus, both FT-NIR spectroscopy and multispectral image analysis proved to be promising tools for classifying cotton seed lots according to their physiological potential.

Keywords: *Gossypium hirsutum* L., seeds vigor, classification models, machine learning, artificial intelligence.

INTRODUÇÃO

O algodão é uma das principais commodities mundiais, destacando-se pelo uso de tecnologias sofisticadas nos sistemas de produção (ABRAPA, 2025). Dessa forma, sementes com alta qualidade fisiológica constituem o insumo básico para o estabelecimento da cultura, contribuindo diretamente para o crescimento das plantas e a produtividade (Krzyzanowski e Ohlson, 2024). A obtenção de sementes de algodão de alta qualidade requer o uso de tecnologias específicas tanto no campo, como reguladores de crescimento vegetal e colheita mecanizada, quanto no beneficiamento pós-colheita, que envolve etapas como o descaroçamento e o deslintamento, as quais podem causar danos às sementes (Krzyzanowski e Ohlson, 2024). Além disso, por se tratar de sementes oleaginosas, estas são mais suscetíveis à deterioração (Dowd et al., 2019).

Assim, o uso de tecnologias modernas para o controle de qualidade de lotes de sementes ao longo de todas essas etapas de produção é fundamental. Métodos rápidos, automatizados, eficientes e não destrutivos, que permitam uma tomada de decisão mais ágil quanto ao manejo de lotes nas empresas de sementes, tornam-se, portanto, essenciais. Nesse contexto, a aplicação de tecnologias modernas aliadas a algoritmos de inteligência artificial pode ser uma alternativa eficiente para estimar o potencial fisiológico das sementes, podendo posteriormente ser complementada por métodos tradicionais (Xia et al., 2019). Entre essas tecnologias, destacam-se a espectroscopia no infravermelho próximo (NIR) e a análise de imagens multiespectrais.

A espectroscopia no infravermelho próximo (NIR) baseia-se na absorção da radiação eletromagnética em comprimentos de onda que variam de 780 a 2.500 nm (Xia et al., 2019). Quando a radiação eletromagnética interage com uma amostra, ocorre interação com grupos funcionais formados por ligações químicas como C–H, O–H, S–H e N–H, permitindo que as leituras espectrais sejam associadas à presença de água e de compostos orgânicos, como carboidratos, proteínas e lipídios (Larios et al., 2020; Kumar et al., 2015). Avanços recentes nessa técnica confirmam sua aplicabilidade em estudos que estimam o potencial fisiológico de sementes de algodão (Mayrinck et al., 2020), soja (Soares et al., 2024), milho (Schütz et al., 2022), lentilha (Limão et al., 2025), grão-de-bico (Ribeiro et al., 2021), *Brachiaria* (Medeiros et al., 2020; Souza et al., 2023) e trigo (Xia et al., 2019), bem como na distinção entre sementes de algodão convencionais e geneticamente modificadas (Mata et al., 2022).

A análise de imagens multiespectrais, por sua vez, baseia-se nas propriedades eletromagnéticas dos tecidos das sementes, especialmente na reflectância e/ou fluorescência, e envolve a captura de imagens com a exposição das sementes à luz em diferentes comprimentos

de onda, geralmente na faixa de 400 a 1.000 nm (ElMasry et al., 2020). O princípio dessa técnica reside no fato de que cada semente possui características intrínsecas de estrutura celular e composição química; assim, a luz absorvida e refletida gera variações espectrais que podem ser associadas aos atributos de interesse (Shrestha et al., 2015; ElMasry et al., 2020). Essa técnica foi utilizada para a identificação de cultivares de tomate (Shrestha et al., 2015) e trigo (Vresak et al., 2016), distinção entre sementes de arroz transgênicas e não transgênicas (Liu et al., 2014), classificação de lotes de sementes de soja e avaliação da integridade física (Medeiros et al., 2020; Silva et al., 2021), bem como para a estimativa da qualidade fisiológica de sementes (Hemender et al., 2018).

Entretanto, a eficiência dessas técnicas é potencializada quando combinadas com métodos adequados de pré-processamento dos dados espectrais e com algoritmos de classificação (Soares et al., 2024). A aplicação de aprendizado de máquina para a interpretação desses dados permite o desenvolvimento de modelos preditivos precisos, capazes de estimar o potencial fisiológico das sementes com base em suas assinaturas espectrais. Os métodos de pré-processamento são importantes para reduzir ruídos, bem como efeitos aditivos e multiplicativos nos espectros, possibilitando a obtenção de informações reais sobre o material em estudo. Entre os métodos de pré-processamento mais comuns destacam-se as correções de espalhamento (*Standard Normal Variate* – SNV e *Multiplicative Scatter Correction* – MSC) e as derivadas (1ª ou 2ª derivada de Savitzky-Golay – SG) (Dhanoa et al., 1994; Rinnan et al., 2009).

Algoritmos como a Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA), *Random Forest* e Redes Neurais constituem alternativas eficazes de modelagem, capazes de reconhecer padrões nos comprimentos de onda de dados espectrais e multiespectrais (Li et al., 2019). Em estudos de qualidade de sementes, pesquisadores têm demonstrado a eficácia de modelos PLS-DA na diferenciação de lotes de sementes com base no potencial fisiológico de lentilha (Limão et al., 2025), soja (Soares et al., 2024; Silva et al., 2024) e *Brachiaria* (Souza et al., 2023), bem como na identificação de diferenças na qualidade de sementes de grão-de-bico após a aplicação de herbicidas em pré-colheita (Ribeiro et al., 2021), entre outros. O algoritmo *Random Forest* tem sido utilizado na classificação da qualidade de sementes de melancia (Liu et al., 2019) e crambe (Medeiros et al., 2021), na avaliação da viabilidade (Baek et al., 2019) e na identificação de níveis de deterioração em sementes de soja (Silva et al., 2021). De forma semelhante, Redes Neurais têm sido aplicadas para classificar sementes de soja (Soares et al., 2023) e feijão (ElMasry et al., 2019) com base em níveis de deterioração,

identificar cultivares de tomate (Shrestha et al., 2015) e detectar padrões de maturação de sementes de soja (Batista et al., 2022).

Nesse contexto, o presente estudo teve como objetivo avaliar a eficiência da análise de imagens multiespectrais e da espectroscopia no infravermelho próximo (NIR), combinadas com algoritmos de aprendizado de máquina, na classificação de sementes de algodão de acordo com seu potencial fisiológico.

MATERIAIS E MÉTODOS

A pesquisa foi conduzida no Laboratório de Pesquisa em Sementes e no Laboratório de Agroenergia do Departamento de Agronomia da Universidade Federal de Viçosa, Viçosa, MG, Brasil. Foram utilizados oito lotes comerciais de sementes da cultivar 21065TLP. As sementes de cada lote foram inicialmente avaliadas quanto ao grau de umidade utilizando o método da estufa a 105°C, por 24 horas, segundo as Regras para Análise de Sementes (Brasil, 2025).

Em seguida, as sementes de cada lote foram submetidas à análise por espectroscopia no infravermelho próximo (FT-NIR) e à aquisição de imagens multiespectrais.

Espectroscopia no Infravermelho Próximo e Análise de Imagens Multiespectrais

Aquisição dos espectros FT-NIR: Para cada lote, 200 sementes foram selecionadas aleatoriamente, totalizando 1.600 sementes avaliadas individualmente. Os dados espectrais de cada semente foram obtidos utilizando um espectrômetro de infravermelho próximo por transformada de Fourier (FT-NIR) (Thermo Scientific Antaris II). Os espectros foram expressos em absorbância ($\log 1/R$). Cada espectro individual foi composto por 3.111 bandas na faixa de comprimento de onda de 1.000 a 2.500 nm. Foram realizadas trinta varreduras sucessivas, sendo calculado o espectro médio de cada semente.

Análise de imagens multiespectrais: As mesmas sementes utilizadas na análise por NIR foram submetidas à aquisição de imagens multiespectrais. Para a captura das imagens, foi utilizado um protótipo equipado com câmera digital e LEDs (Viçosa, Minas Gerais, Brasil), que forneceu imagens UV-VIS-NIR nos comprimentos de onda de 395, 460, 520, 585, 620, 740, 850 e 940 nm, totalizando oito bandas espectrais. As imagens foram capturadas e salvas no formato TIFF (Brochado et al., 2024).

Para a análise das imagens, os níveis de cinza de cada semente foram extraídos de cada banda espectral utilizando o software ImageJ (ImageJ2 Version, Bethesda, Maryland, EUA). As imagens foram processadas para a seleção da região de interesse por meio do método de

Otsu (Otsu, 1979), sendo obtidos os valores médios de escala de cinza de cada semente individual em cada banda espectral.

Os dados obtidos por espectroscopia NIR e análise de imagens foram pré-processados e utilizados para o treinamento dos modelos de classificação por meio do pacote *Caret* do software R (Kuhn, 2008).

Após a aquisição dos espectros NIR e das imagens multiespectrais, as mesmas sementes de cada lote foram submetidas ao teste de germinação.

Para tanto, quatro repetições de 50 sementes foram dispostas sobre papel toalha umedecido com água em quantidade equivalente a 2,5 vezes o peso do papel seco. Os rolos foram preparados e mantidos em germinador a 25 °C. Foram realizadas contagens aos quatro e doze dias após a semeadura, calculando-se a porcentagem de plântulas normais (Brasil, 2025).

Delineamento Experimental e Análise Estatística

Teste de Germinação: O experimento foi conduzido em delineamento inteiramente casualizado, com oito lotes de sementes e quatro repetições. Os dados foram submetidos ao teste de normalidade dos erros (Shapiro-Wilk) e ao teste de homogeneidade de variâncias (Bartlett). Em seguida, realizou-se a análise de variância (ANOVA) e as médias obtidas para cada lote foram comparadas pelo teste de Tukey ($p < 0,05$). As análises estatísticas foram realizadas no software R, versão 4.3.3 (R Core Team, 2025).

Espectroscopia no Infravermelho Próximo e Análise de Imagens Multiespectrais

Algoritmos de Pré-processamento: Os dados espectrais NIR originais foram pré-processados utilizando os seguintes métodos: *Standard Normal Variate* (SNV); *Multiplicative Scatter Correction* (MSC); derivadas de Savitzky-Golay de 1ª e 2ª ordem, com janela de 7 pontos. Para os dados multiespectrais (valores de cinza), foram utilizados apenas os métodos de pré-processamento SNV e MSC.

Algoritmos de Classificação: Três modelos de classificação foram aplicados aos espectros obtidos com diferentes técnicas de pré-processamento e aos valores de cinza extraídos das imagens multiespectrais: Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA), *Random Forest* (RF) e Redes Neurais (NN). Para os algoritmos PLS-DA, *Random Forest* e Redes Neurais, foi utilizado o pacote *Caret* (Kuhn, 2008). O algoritmo *Random Forest* foi executado com 200 árvores de decisão, enquanto a Rede Neural utilizou 20 neurônios na camada oculta.

Desenvolvimento dos Modelos de Classificação: Os lotes de sementes foram classificados em três categorias com base nos resultados do teste de germinação: alto potencial fisiológico (germinação $\geq 91\%$), médio potencial fisiológico (germinação entre 85–90%) e baixo potencial fisiológico (germinação $\leq 84\%$). Assim, a classe de alto potencial foi composta por 400 sementes, a de médio potencial por 800 sementes e a de baixo potencial por 400 sementes. Para a análise de imagens multiespectrais, o banco foi de 100 sementes por lote; dessa forma, a classe de alto potencial foi composta por 200 sementes, a de médio potencial por 400 sementes e a de baixo potencial por 200 sementes.

Validação dos Modelos: Os modelos foram desenvolvidos utilizando 70% dos dados para treinamento e 30% para validação. Os modelos PLS-DA, *Random Forest* e Redes Neurais foram avaliados por meio das métricas de acurácia (Eq. 1), coeficiente kappa (κ) (Eq. 2), sensibilidade e especificidade, conforme as seguintes equações:

$$Accuracia = \frac{Corretas}{total} \quad (1)$$

$$K = \frac{Po - Pe}{1 - Pe} \quad (2)$$

em que Po é a proporção de concordância observada e Pe é a proporção de concordância esperada.

$$Sensibilidade = \frac{TP}{TP + FN}$$

$$Especificidade = \frac{TN}{TN + FP}$$

Em que TP = verdadeiro positivo; TN = verdadeiro negativo; FN = falso negativo; FP= falso positivo.

RESULTADOS

Grau de umidade e Germinação

O grau de umidade das sementes apresentou uniformidade, variando de 9,4 a 10,3% (Tabela 1). Trata-se de um fator importante para as análises espectrais, uma vez que o conteúdo de água nas sementes pode interferir na absorção da luz, em especial no comprimento de onda do infravermelho (Osborne, Fearn e Hindle, 1993).

Observou-se que os lotes 1 e 5 apresentaram os maiores valores de germinação seguidos pelos lotes 2, 3, 6 e 7, enquanto os lotes 4 e 8 exibiram os menores valores. Portanto, com base nestes resultados foram observadas três classes de sementes quanto ao potencial fisiológico.

Tabela 1. Grau de umidade e germinação de oito lotes de sementes da cultivar de algodão 21065TLP.

| Lotes | Grau de umidade (%) | Germinação (%) |
|--------|---------------------|----------------|
| 1 | 10,1 | 96 a |
| 2 | 9,8 | 90 b |
| 3 | 9,4 | 89 b |
| 4 | 9,4 | 82 c |
| 5 | 10,3 | 95 a |
| 6 | 9,5 | 89 b |
| 7 | 9,7 | 87 b |
| 8 | 9,7 | 79 c |
| CV (%) | - | 1,86 |

CV = coeficiente de variação. Médias seguidas pela mesma letra nas colunas não diferem entre si pelo teste de Tukey ($p < 0,05$).

Espectroscopia no Infravermelho Próximo (FT-NIR)

1.600 espectros brutos e a média dos espectros brutos é apresentada na Figura 1(A) e 1(B), de acordo com as classes de potencial fisiológico: alto (germinação $\geq 91\%$), médio (germinação entre 85–90%) e baixo potencial fisiológico (germinação $\leq 84\%$).

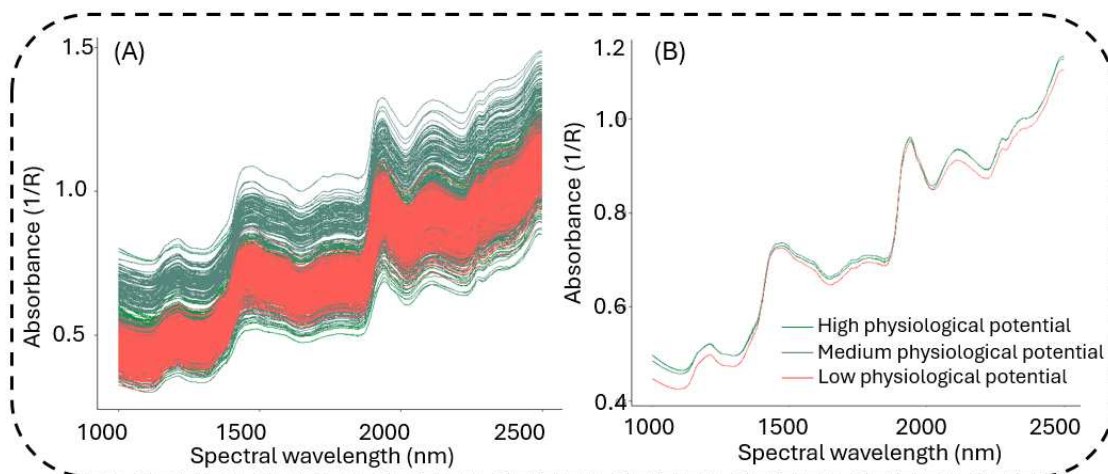


Figura 1. Espectros brutos (A) e espectros médios (B) obtidos por meio da técnica FT-NIR para sementes de oito lotes de algodão, de acordo com seus níveis de potencial fisiológico.

Os modelos de classificação testados apresentaram elevado desempenho na etapa de validação (Tabela 2). Os classificadores PLS-DA, *Random Forest* e Redes Neurais obtiveram menores valores de acurácia e coeficiente kappa quando os dados foram pré-processados utilizando SNV e MSC. No entanto, os modelos PLS-DA, *Random Forest* e Redes Neurais apresentaram desempenho superior quando aplicada a 2ª derivada de Savitzky-Golay, alcançando valores de acurácia na validação de 98%, 93% e 51%, respectivamente (Tabela 2). Em relação ao coeficiente kappa, os resultados atingiram 96% para o PLS-DA, 89% para o *Random Forest* e 93% para as Redes Neurais. Observa-se também valores superior a 90% para a métrica de especificidade para todos os modelos. Entretanto, apenas o modelo PLS-DA e Redes Neurais apresentaram sensibilidade superior a 90% (Tabela 2).

Tabela 2. Acurácia e coeficiente kappa dos modelos de classificação PLS-DA, *Random Forest* e Redes Neurais para diferentes métodos de pré-processamento de oito lotes de sementes de algodão, utilizando dados espectrais obtidos por FT-NIR.

| Modelo | Pré-processamento | Validação | | | |
|---------|----------------------|-----------|-------|---------------|----------------|
| | | Acurácia | Kappa | Sensibilidade | Especificidade |
| (n=480) | Raw | 0,80 | 0,66 | 0,76 | 0,87 |
| | PLS- SNV | 0,81 | 0,69 | 0,77 | 0,88 |
| | DA MSC | 0,80 | 0,66 | 0,75 | 0,87 |
| | 1st Derivative of SG | 0,88 | 0,79 | 0,84 | 0,91 |
| | 2nd Derivative of SG | 0,98 | 0,96 | 0,93 | 0,96 |
| (n=480) | Raw | 0,71 | 0,51 | 0,66 | 0,82 |
| | Random SNV | 0,89 | 0,82 | 0,87 | 0,93 |
| | Forest MSC | 0,89 | 0,82 | 0,87 | 0,93 |
| | 1st Derivative of SG | 0,91 | 0,85 | 0,91 | 0,95 |
| | 2nd Derivative of SG | 0,93 | 0,89 | 0,88 | 0,94 |
| (n=480) | Raw | 0,74 | 0,59 | 0,70 | 0,81 |
| | Neural SNV | 0,91 | 0,86 | 0,89 | 0,94 |
| | Network MSC | 0,91 | 0,86 | 0,90 | 0,94 |
| | 1st Derivative of SG | 0,94 | 0,90 | 0,91 | 0,95 |
| | 2nd Derivative of SG | 0,96 | 0,93 | 0,93 | 0,96 |

n = número de espectros utilizados para validação. SNV = *Standard Normal Variate*; MSC = *Multiplicative Scatter Correction*; SG = *Savitzky-Golay*.

A Figura 2 apresenta a matriz de confusão obtida a partir dos resultados de validação do método de pré-processamento com a 2ª derivada de Savitzky–Golay para os modelos PLS-DA, *Random Forest* e Redes Neurais. De modo geral, observa-se elevada acurácia de classificação para as classes de alto, médio e baixo potencial fisiológico em todos os modelos testados, com mínima distribuição de erros entre as classes. No entanto, o modelo PLS-DA alcançou a maior proporção de acerto, com 100% para a classe de alto, 98% para a classe de médio e 97% para a classe de baixo potencial fisiológico. As Redes Neurais, por sua vez, atingiram 100% de proporção de acerto para a classe de alto potencial fisiológico, mas apresentaram maior distribuição de erros nas classes de médio (95%) e baixo (94%) potencial. O modelo *Random Forest*, por outro lado, exibiu maior distribuição de erros entre as classes em comparação aos demais algoritmos, com proporção de acerto de 92% para a classe de alto, 94% para a classe de médio e 93% para a classe de baixo potencial fisiológico.

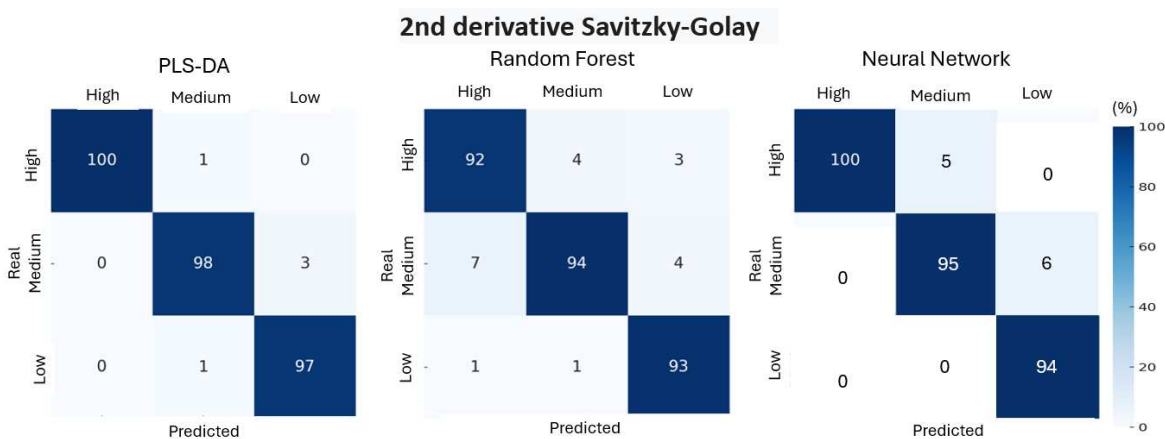


Figura 2. Matriz de confusão dos dados de validação para a classificação dos níveis de potencial fisiológico de lotes de sementes de algodão utilizando os modelos PLS-DA, *Random Forest* e Redes Neurais, com pré-processamento pela 2ª derivada de Savitzky-Golay.

A importância das variáveis, destacando os comprimentos de onda que mais contribuíram para o desenvolvimento do modelo PLS-DA na classificação dos três níveis de potencial fisiológico de sementes de algodão foi apresentada na Figura 3. O modelo PLS-DA foi escolhido por apresentar maiores valores de acurácia e coeficiente kappa quando aplicada a

2ª derivada de Savitzky-Golay. Os comprimentos de onda mais importantes situaram-se nas faixas de 1.000 nm, 1.300–1.450 nm, 1.700–1.900 nm e 2.250–2.400 nm (Figura 3).

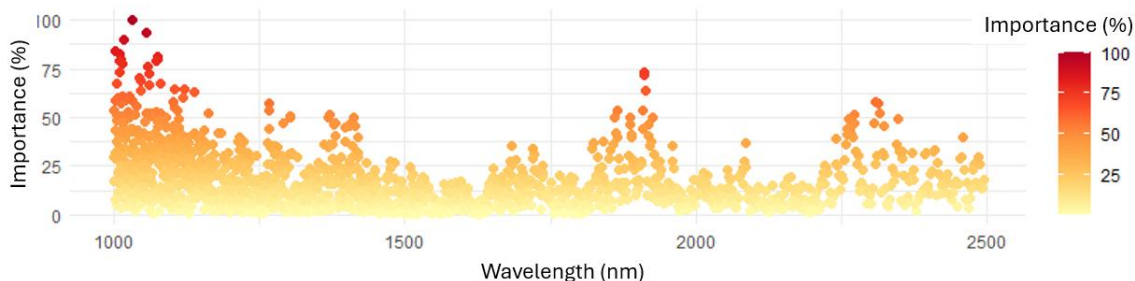


Figura 3. Importância das variáveis de comprimento de onda que mais contribuíram para o desenvolvimento do modelo PLS-DA, de acordo com os níveis de potencial fisiológico de sementes de algodão cv. 21065TLP, utilizando a técnica NIR.

Análise de imagens multiespectrais

Observam-se diferenças no padrão dos valores de níveis de cinza obtidos para os diferentes níveis de potencial fisiológico das sementes: alto (germinação $\geq 91\%$), médio (germinação entre 85–90%) e baixo potencial fisiológico ($< 84\%$) (Figura 4A, B).

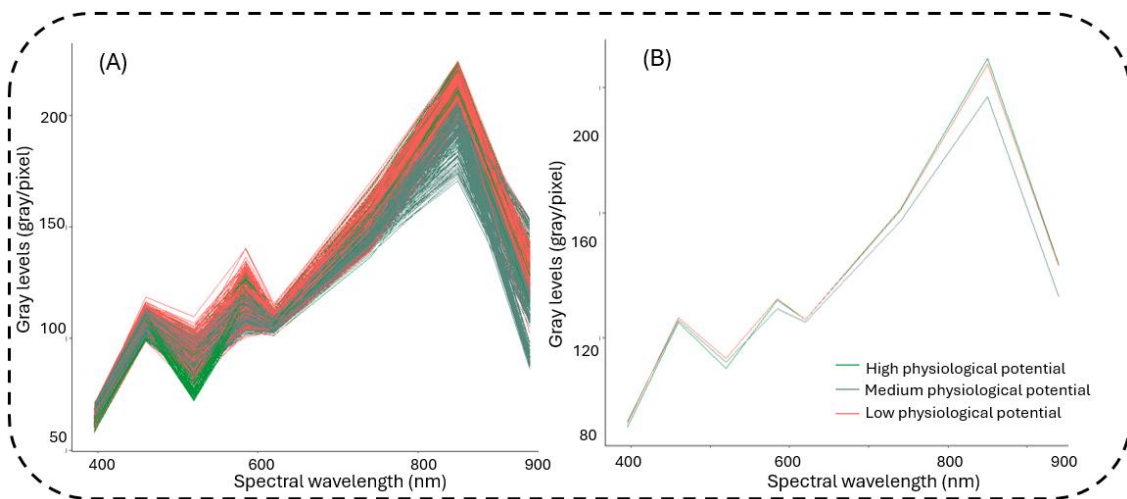


Figura 4. Valores brutos de níveis de cinza (A) e médias dos valores de níveis de cinza (B) obtidos a partir da análise de imagens multiespectrais de oito lotes de sementes de algodão, de acordo com os níveis de potencial fisiológico.

Os resultados dos modelos de classificação para os dados da análise de imagens multiespectrais são apresentados na Tabela 3. Na validação, os modelos PLS-DA, *Random Forest* e Redes Neurais apresentaram melhor desempenho quando nenhum pré-processamento

foi aplicado, ou seja, utilizando os dados brutos. Ao comparar os modelos PLS-DA, *Random Forest* e Redes Neurais, o modelo *Random Forest* apresentou o melhor desempenho com os dados brutos, alcançando acurácia de 81% e coeficiente kappa de 71%, seguido pelo pré-processamento MSC, que apresentou resultados semelhantes, com acurácia de 81% e kappa de 70%. O modelo de Redes Neurais apresentou desempenho intermediário, com acurácia de 80% e kappa de 69% utilizando os dados brutos, seguido do pré-processamento MSC (71% de acurácia e 56% de kappa). No entanto, o modelo de classificação PLS-DA apresentou o menor desempenho, com os dados brutos atingindo 72% de acurácia, enquanto os demais métodos de pré-processamento ficaram abaixo de 70%, indicando menor ajuste do modelo. De forma semelhante, para a métrica kappa, todos os métodos de pré-processamento apresentaram variações entre 48% e 56%.

Tabela 3. Valores de acurácia e coeficiente kappa dos modelos de classificação PLS-DA, *Random Forest* e Redes Neurais para os métodos de pré-processamento SNV e MSC de oito lotes de sementes de algodão, na validação, utilizando dados de análise de imagens multiespectrais.

| Modelo | Pré-processamento | Validação | | | |
|------------------------------|-------------------|-----------|-------|---------------|----------------|
| | | Acurácia | Kappa | Sensibilidade | Especificidade |
| PLS-DA (n=480) | Raw | 0,72 | 0,56 | 0,69 | 0,84 |
| | SNV | 0,67 | 0,48 | 0,63 | 0,81 |
| | MSC | 0,67 | 0,48 | 0,63 | 0,81 |
| Random Forest (n=480) | Raw | 0,81 | 0,71 | 0,81 | 0,89 |
| | SNV | 0,80 | 0,69 | 0,78 | 0,89 |
| | MSC | 0,81 | 0,70 | 0,79 | 0,89 |
| Neural Network (n=480) | Raw | 0,80 | 0,69 | 0,90 | 0,94 |
| | SNV | 0,76 | 0,62 | 0,73 | 0,86 |
| | MSC | 0,71 | 0,56 | 0,70 | 0,84 |

n = número de espectros utilizados para validação. SNV = *Standard Normal Variate*; MSC = *Multiplicative Scatter Correction*.

A Figura 5 apresenta os resultados da matriz de confusão obtidos na validação utilizando os dados brutos para os três algoritmos testados (PLS-DA, *Random Forest* e Redes Neurais).

Observou-se menor distribuição de erros para o modelo *Random Forest*, conforme Tabela 3. Os modelos de Redes Neurais e PLS-DA apresentaram menor eficiência quando comparados ao modelo *Random Forest* (Figura 5).

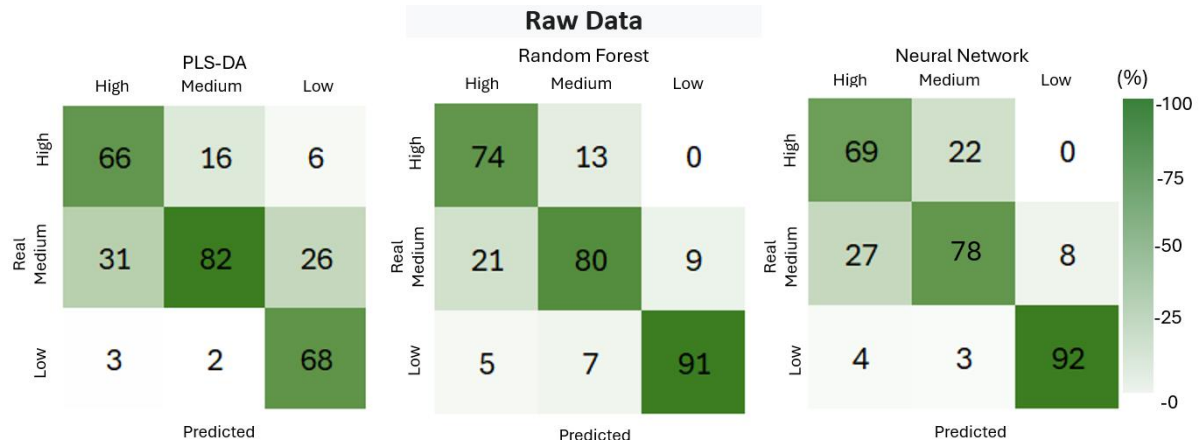


Figura 5. Matriz de confusão dos dados de validação para a classificação de lotes de sementes de algodão de acordo com seu nível de potencial fisiológico, utilizando os modelos PLS-DA, *Random Forest* e Redes Neurais.

A Figura 6 ilustra o gráfico de importância das variáveis, evidenciando os comprimentos de onda que mais contribuíram para o desenvolvimento do modelo de classificação *Random Forest* para as três classes de potencial fisiológico de sementes de algodão. As faixas espectrais de 480 nm, 510 nm, 620 nm e 740 nm foram as mais importantes para o desenvolvimento do modelo de classificação (Figura 6).

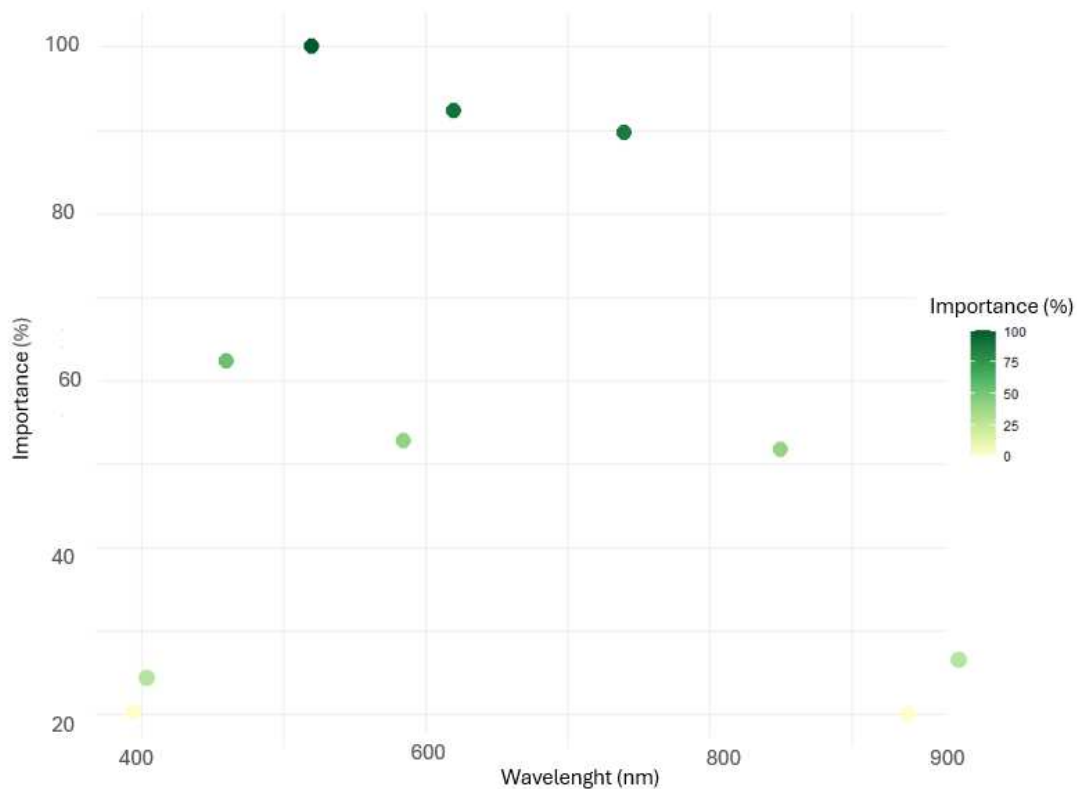


Figura 6. Importância das variáveis de comprimento de onda que mais contribuíram para o desenvolvimento do modelo *Random Forest*, de acordo com os níveis de potencial fisiológico de sementes de algodão cv. 21065TLP, utilizando a análise de imagens multiespectrais.

DISCUSSÃO

Caracterização do Potencial Fisiológico

O grau de umidade das sementes foi semelhante entre os lotes variando entre 9,4 e 10,3% (Tabela 1), sendo esta uma característica importante a ser considerada durante a realização dos testes de qualidade fisiológica para que as avaliações possam ser padronizadas com resultados consistentes (Marcos-Filho, 2020). Pelos testes de avaliação da qualidade fisiológica (Tabela 1), foi possível observar diferenças entre o potencial fisiológico dos lotes. Em geral, os lotes 1 e 5 tiveram melhor desempenho tanto no teste de germinação quanto nos testes de vigor, principalmente em relação aos lotes 4 e 8, com pior desempenho.

O teste de germinação é realizado em condições ideais, permitindo que as sementes expressem seu máximo potencial fisiológico, enquanto a primeira contagem de germinação indica a velocidade de germinação (Brasil, 2025). Já os testes de envelhecimento acelerado e germinação em baixa temperatura avaliam o comportamento das sementes sob condições de

estresse (Dias e Marcos-Filho, 2020; Marcos-Filho, 2020a), de modo que sementes com maior vigor apresentam melhor desempenho do que sementes com menor vigor sob amplas condições de ambiente (Marcos-Filho 2020b). A utilização de mais de um teste para avaliação do potencial fisiológico das sementes é importante, pois o desempenho do lote pode variar em função do princípio do teste. O lote 8, por exemplo, mostrou-se bastante sensível às condições de estresse do teste de envelhecimento acelerado, com valor de 32% de plântulas normais, sendo inferior aos demais, o que não ocorreu no teste de germinação a baixa temperatura, no qual o lote atingiu valor de 80%, não diferindo dos lotes 2, 3, 6 e 7. Assim, classificar os lotes quanto ao potencial fisiológico por métodos tradicionais traz informações relevantes para o controle de qualidade interno nas empresas de sementes (Marcos-Filho 2020b).

Contudo, na atualidade, há uma busca por técnicas que exigem menor tempo para a obtenção de resultados e, se possível, não destrutivos e automatizados permitindo agilizar as tomadas de decisão quanto ao destino dos lotes produzidos a cada safra. Neste contexto, entre estas tecnologias mais recentes, a análise de imagens multiespectrais e a espectroscopia NIR vêm sendo aplicadas para várias espécies como amendoim (Oliveira et al., 2024), soja (Soares et al. 2024, Larios et al. 2020, Batista et al. 2022, Petronilio et al. 2025), milho (Schütz et al. 2022), tomate (Shrestha et al. 2016), lentilha (Limão et al., 2025) e braquiária (Medeiros et al. 2020, Souza et al. 2023).

Esses métodos são particularmente interessantes para sementes de algodão, cuja qualidade pode ser influenciada negativamente por fatores como condições de produção em campo, estágio de maturação e manejo durante a colheita, manejo e processamento pós-colheita, incluindo descaroçamento e deslintamento químico, e condições de armazenamento (Snipes e Baskin 1994). Esses fatores podem acelerar o processo de deterioração e, conseqüentemente, causar alterações bioquímicas que se refletem na germinação e no estabelecimento do algodoeiro no campo, o que pode impactar a produtividade da cultura (Krzyzanowski e Ohlson 2024).

Espectroscopia no infravermelho próximo (NIR)

A espectroscopia no infravermelho próximo (NIR) baseia-se no princípio da interação da radiação eletromagnética da região do infravermelho próximo (800 a 2500 nm) com as moléculas presentes na semente (Mayrinck et al. 2020). A radiação é absorvida principalmente por ligações químicas do tipo C-H, O-H, N-H e S-H, que são associadas a compostos orgânicos como água, proteínas, lipídeos e carboidratos (Koczoń et al. 2023), resultando em padrões

espectrais distintos. Assim, considerando que o processo de deterioração provoca alterações bioquímicas e estruturais nas sementes como degradação de lipídeos, oxidação de proteínas e outras moléculas, especialmente em espécies oleaginosas como o algodão, esta técnica pode ser promissora como ferramenta para a classificação de lotes quanto ao potencial fisiológico (Masilamani et al. 2020).

No nosso estudo, questionamos se a técnica FT-NIR seria eficiente para classificar lotes de algodão com diferentes níveis de potencial fisiológico, bem como buscamos definir métodos adequados para o pré-processamento dos dados espectrais e selecionar modelos de classificação. Com o pré-processamento é possível eliminar ou amenizar variâncias indesejáveis intrínsecas ao equipamento ou às amostras (Sena e Almeida, 2018). A escolha de um método inadequado pode resultar em perda de informações importantes (Pasquini, 2018), o que pode comprometer a definição do modelo de classificação final. Assim, a aplicação de algoritmos de aprendizado de máquina para modelar a relação entre os dados espectrais e a composição química das sementes (Xia et al. 2019), que é alterada de acordo com o seu nível de qualidade (Souza et al., 2023), é uma etapa fundamental para o sucesso do uso da técnica FT-NIR.

Dentre os modelos testados, o PLS-DA e o pré-processamento da segunda derivada de SG mostraram-se mais eficientes com valores de acurácia de 98% e kappa de 96% (Tabela 2), os quais foram superiores aos obtidos com os demais modelos. Este algoritmo vem sendo utilizado com sucesso na construção de modelos preditivos que buscam classificar sementes quanto ao potencial fisiológico com base nos espectros NIR (Soares et al. 2024, Silva et al. 2024, Limão et al. 2025, Masilamani et al. 2020, Kusumaningrum et al. 2018). O algoritmo Savitzky-Golay é um algoritmo de suavização espectral. Ele não só ajuda a suavizar o sinal, removendo o ruído de alta frequência, mas também permite o cálculo das derivadas de primeira e segunda ordem úteis para detectar mudanças rápidas e curvaturas nos dados espectrais (Rinnan et al. 2009). Neste estudo, a segunda derivada de Savitzky-Golay mostrou-se mais eficiente como modelo de classificação quanto ao potencial fisiológico. Essa derivada remove a linha de base e a tendência linear (Rinnan et al. 2009).

Os modelos foram validados por meio das métricas de acurácia e kappa conforme observado na Tabela 3. A acurácia representa a proporção de previsões corretas em relação ao total de previsões realizadas (Han et al., 2012). Já o coeficiente Kappa de Cohen é utilizado para quantificar o grau de concordância. Esse coeficiente representa uma medida de concordância ajustada, sendo calculado a partir da proporção de concordância observada e da

concordância esperada (Cohen, 1960). Dessa forma, o Kappa fornece uma estimativa mais robusta da consistência entre avaliadores ou sistemas de classificação do que a simples porcentagem de acertos, especialmente em situações em que as categorias apresentam distribuições desbalanceadas. Segundo Cohen (Cohen, 1960), valores de kappa próximo de 1 indica concordância perfeita. Logo, no nosso estudo obtivemos valores de kappa de 96%, ou seja, alta concordância dos modelos quando se aplicou a 2ª derivada SG.

O desempenho das técnicas de aprendizado de máquina PLS-DA, *Random Forest*, e Redes Neurais foi avaliado por meio da análise da matriz de confusão (Figura 2). Os resultados indicam alta acurácia de todos os modelos utilizados para estimar a classificação dos lotes de alto, médio e baixo potencial fisiológico. Contudo, é importante considerar os resultados específicos de cada algoritmo para uma análise detalhada da performance de cada um. Os menores erros na distribuição das classes ocorreram quando se utilizou o modelo PLS-DA, com o qual 100% das predições foram corretas para a classe de alto potencial, enquanto apenas 2% dos dados da classe de médio e 3% da classe de baixo potencial foram classificados incorretamente. Portanto, foi o modelo com menor distribuição de erros entre as classes, em comparação com o *Random Forest* e Redes Neurais. O PLS-DA foi desenvolvido para utilizar a calibração multivariada e aplicado para reduzir a dimensão do banco de dados, sendo aplicados para fins de seleção e classificação de amostras, determinando a qual classe a amostra desconhecida pode pertencer (Ruiz-Perez et al. 2020, Santana et al. 2020). Alguns estudos com dados espectrais já comprovaram a aplicação desse algoritmo em sementes de amendoim (Petronilio et al. 2025), grão-de-bico (Ribeiro et al. 2021), braquiária (Souza et al. 2023, Medeiros et al. 2020), soja (Soares et al. 2024, Silva et al. 2024) e milho (Andrade et al. 2020).

As bandas espectrais que tiveram maior contribuição para o desenvolvimento do modelo PLS-DA foram 1,000 nm, 1.300–1.450 nm, 1.700–1.900 nm, e 2.250–2.400 nm (Figura 3). No comprimento de onda em torno de 1.000 nm, predominam os grupos hidroxila (O-H) associados aos carboidratos (Workman e Weyer 2007). Já a faixa entre 1.300 e 1.450 nm está associada a grupos funcionais carboxílicos e hidroxílicos relacionados a ácidos graxos monoinsaturados, poli-insaturados e saturados (C-H) (Ozaki et al. 2006). Esse intervalo também está associado a proteínas (O-H) (Ambrose et al. 2016, Xu et al. 2020).

Nos intervalos espectrais de 1.800 a 1.900 nm, também houve contribuição significativa para o desenvolvimento do modelo. O comprimento de onda de 1.860 nm está relacionado a proteínas (Kusumaningrum et al. 2018), enquanto o intervalo espectral de 1.930 nm está associado a grupos O-H, carboidratos e água (Soares et al., 2016). De acordo com Silva et al.

(2024), esses intervalos espectrais também foram cruciais para o desenvolvimento do modelo de classificação do vigor de sementes de soja utilizando espectroscopia no infravermelho próximo. Os comprimentos de onda de 2.250 a 2.500 nm estão relacionados aos teores de lipídios e proteínas (Xu et al. 2020, Ambrose et al. 2016). O intervalo entre 2.308 e 2.348 nm representa absorção de lipídios (Ozaki et al. 2006). O comprimento de onda de 2.308 nm também pode ser usado para diferenciar ácidos graxos monoinsaturados de poli-insaturados (Hourant et al. 2000). Já 2.344 nm está associado ao grupo funcional C-H e está relacionado ao teor de carboidratos (celulose) (Soares et al., 2016).

Portanto, as diferenças observadas nos espectros de acordo com as classes de potencial fisiológico refletem mudanças na composição química das sementes, especialmente lipídios, carboidratos e proteínas. Estes compostos constituem as principais reservas das sementes de algodão, que são classificadas como oleaginosas, o que as torna mais propensas à deterioração durante as diferentes etapas do processo produtivo (Mata et al., 2022; Soares et al., 2016). Portanto, a intensidade das alterações bioquímicas decorrentes da deterioração, conseqüentemente, tem reflexos sobre o potencial fisiológico das sementes (Marcos-Filho, 2015).

Análise de imagens multiespectrais

Atualmente existe grande demanda por ferramentas rápidas, automatizadas e precisas para o controle de qualidade de sementes. Neste contexto, o uso de imagens multiespectrais associadas com algoritmos de aprendizado de máquina surge como uma alternativa promissora nesse segmento (Petronílio et al., 2025). Da mesma forma que o NIR, questionamos se a tecnologia de imagens multiespectrais seria eficiente para classificar os lotes de sementes de algodão com diferentes níveis de potencial fisiológico.

A análise de imagens multiespectrais tem como princípio a captura de imagens das amostras em diferentes faixas do espectro eletromagnético, geralmente do visível ao infravermelho próximo, explorando o fato de que cada material possui padrões específicos de absorção, transmissão e reflexão da radiação eletromagnética nos diferentes comprimentos de onda (ElMasry et al. 2019). Assim, ao combinar as informações espectrais e morfológicas (cor, brilho, textura) torna-se possível identificar diferenças relacionadas à composição físico-química e estrutural das amostras (ElMasry et al. 2019). Em sementes, por exemplo, essa ferramenta emergente possibilita detectar variações associadas ao conteúdo de lipídeos, proteínas, carboidratos e água decorrentes do processo de deterioração, possibilitando alto

potencial de automação na determinação do potencial fisiológico (ElMasry et al. 2019, Petronilio et al. 2025).

Neste estudo, fizemos o imageamento multiespectrais e separamos os lotes em três classes de potencial fisiológico (alto, médio e baixo). Foi possível observar diferenças das classes conforme os valores de cinza (Figura 4A) e a média dos valores de cinza (Figura 4B). Em seguida, aplicamos e testamos diferentes algoritmos para desenvolvimento de modelos com base nos níveis de potencial fisiológico estabelecidos.

Algoritmos robustos permitem capturar relações lineares e não lineares e podem alcançar alta acurácia na classificação. Dentre eles, podemos citar o PLS-DA, *Random Forest* e Redes Neurais (Ruiz-Perez et al. 2020, Santana et al. 2020, Braga, Lima-Filho e Ludermir 2007). Entretanto, tais algoritmos operam de modo diferente e podem apresentar diferentes performances (Benedet et al., 2020). Nós obtivemos resultados promissores ao utilizar os dados originais no modelo Random Forest com valores de acurácia de 81% e kappa de 71% na validação (Tabela 3). O Random Forest é um algoritmo de aprendizado de máquina baseado na combinação de múltiplas árvores de decisão para realizar tarefas de classificação. O algoritmo cria classificadores de forma aleatória e faz a combinação de todos os classificadores até chegar no resultado final (Santana et al. 2020). Vale destacar que algoritmos de aprendizado de máquina como Random Forest têm sido cada vez mais utilizados para modelagem e predição do potencial fisiológico de sementes (Medeiros et al. 2020, Silva et al. 2021), contribuindo diretamente para a geração de resultados confiáveis e agilidade na tomada de decisão nas empresas produtoras de sementes.

Conforme observado na Figura 5, o modelo Random Forest apresentou maior quantidade de acertos e menor quantidade de distribuição de erros quando comparado com os demais modelos, sendo 74% para a classe de alta, 80% para médio e 91% para a classe de baixo potencial fisiológico.

Os comprimentos de onda de 480 nm e 510 nm (região azul), conforme mostrado na Figura 6, correspondem a regiões sensíveis a pigmentos naturais no tegumento da semente, como fenóis oxidados, flavonoides e clorofilas residuais, os quais são importantes para diferenciar a qualidade fisiológica, pois alterações químicas durante a deterioração podem afetar a absorção/reflexão nessa faixa de comprimento de onda (Liu et al., 2019). O comprimento de onda de 610 nm (região vermelha) está diretamente relacionado à absorção de clorofila a e b (Liu et al., 2019), 740 nm (região próxima ao vermelho) corresponde à transição entre forte absorção de vermelho e níveis médios de cinza mais altos. Nessa faixa, a reflectância

aumenta à medida que a semente perde clorofila; assim, sementes com maior potencial fisiológico, integridade e viabilidade apresentam valores de reflectância mais altos nessa região (Gitelson et al. 2003). Vale ressaltar também que essa faixa é sensível ao teor de água da semente (Gitelson et al. 2003) (Figura. 6).

As imagens multiespectrais combinada com quimiometria pode fornecer informações ricas e valiosas sobre diversos atributos das sementes, incluindo propriedades morfológicas, físico-químicas e fisiológicas (Petronilio et al., 2025). Essas abordagens têm atraído pesquisadores em todo o mundo, pois oferecem resultados rápidos, precisos e não destrutivos, podendo ser implementadas como ferramentas complementares aos testes tradicionais de avaliação da qualidade de sementes.

O uso e combinação de diferentes técnicas para classificação das sementes quanto ao potencial fisiológico tem aumentado bastante nos últimos anos na área de tecnologia de sementes (Medeiros et al. 2020). No presente estudo nós utilizamos informações obtidas pela técnica do NIR e por um protótipo de equipamento para obtenção de imagens multiespectrais associados a algoritmos de inteligência artificial para obter modelos de classificação. Nossos resultados mostram que ambas as técnicas são ferramentas interessantes para estimar a qualidade de sementes de algodão tendo como base as alterações na sua composição química, levando-se em consideração a alta precisão e acurácia alcançada com os modelos PLS-DA e Random Forest, respectivamente para as análises no NIR e imagens multiespectrais.

É importante considerar que a predição e classificação de sementes de acordo com o potencial fisiológico utilizando espectroscopia no infravermelho próximo e análise de imagens multiespectrais representam um avanço tecnológico significativo para a indústria de sementes, pois permitem resultados rápidos, não destrutivos e precisos. Portanto, esses métodos possuem potencial de aplicação em programas de controle de qualidade de sementes, particularmente para o manejo racional de lotes, evitando custos desnecessários associados ao processamento de lotes de baixo potencial fisiológico e possibilitando tomadas de decisão mais ágeis quanto à destinação dos lotes produzidos a cada safra.

CONCLUSÕES

A espectroscopia no infravermelho próximo (FT-NIR) é eficiente para a classificação de sementes de algodão de acordo com seu potencial fisiológico. Todos os modelos testados mostraram-se promissores, com destaque para o PLS-DA, especialmente quando utilizado o pré-processamento com a 2ª derivada de Savitzky-Golay, alcançando acurácia de até 98%. O

modelo Random Forest, utilizando os dados brutos, foi o mais eficiente na classificação do potencial fisiológico de sementes de algodão por meio de imagens multiespectrais, atingindo acurácia de até 81%.

Esses métodos são promissores para a classificação rápida, não destrutiva e precisa da qualidade de sementes de algodão em programas de controle de qualidade de sementes.

REFERÊNCIAS

ABRAPA. Evolução da balança comercial anual de algodão do Brasil e UFs. 2025b. Disponível em:

<https://app.powerbi.com/view?r=eyJrIjoiMzlmMzNiMjItMjg4Mi00NWYyLWJhYzQtMWVlMmI4OGQ5YjU0IiwidCI6IjRhMDk1OGIzLTg4MWQtNDBmYS05NTU1LTIwODQ1MzdhYWZkMyJ9>. Acesso em: 5 nov. 2025.

AMBROSE, A. et al. High speed measurement of corn seed viability using hyperspectral imaging. *Infrared Physics & Technology*, v. 75, p. 173–179, 2016.

ANDRADE, G. C.; COELHO, C. M. M.; UARROTA, V. G. Modelling the vigour of maize seeds submitted to artificial accelerated ageing based on ATR-FTIR data and chemometric tools. *Heliyon*, v. 6, e03477, 2020.

BATISTA, T. B. et al. A reliable method to recognize soybean seed maturation stages based on autofluorescence-spectral imaging combined with machine learning algorithms. *Frontiers in Plant Science*, v. 13, 2022.

BENEDET, L. et al. Soil subgroup prediction via portable X-ray fluorescence and visible near-infrared spectroscopy. *Geoderma*, v. 365, 114212, 2020.

BRAGA, A. P.; LIMA FILHO, A. C. P.; LUDERMIR, T. B. *Redes neurais artificiais: teoria e aplicações*. 2. ed. Rio de Janeiro: LTC, 2007.

BRASIL. Ministério da Agricultura e Pecuária. Secretaria de Defesa Agropecuária. *Regras para análise de sementes – RAS*. Brasília: MAPA/SDA, 2025. Disponível em: https://wikisda.agricultura.gov.br/pt-br/Laborat%C3%B3rios/Metodologia/Sementes/RAS_2025/cap_15_tol_fiscalizacao. Acesso em: 5 nov. 2025.

BROCHADO, M. G. S. et al. What is the most effective analytical method for quantification and identification of microplastics in contaminated soils? *Environmental Geochemistry and Health*, v. 46, n. 7, p. 260, 2024.

COHEN, J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, v. 20, p. 37–46, 1960.

DHANOVA, M. S.; LISTER, S. J.; SANDERSON, R.; BARNES, R. J. The link between multiplicative scatter correction (MSC) and standard normal variate (SNV) transformations of NIR spectra. *Journal of Near Infrared Spectroscopy*, v. 2, n. 1, p. 43–47, 1994.

DIAS, D. C. F. S. e MARCOS-FILHO, J. Teste de germinação a baixa temperatura. In: ABRATES (Ed.). *Vigor de sementes: conceitos e testes*. Londrina: ABRATES, 2020. p. 317–329.

DOWD, M. K.; MANANDHAR, R.; DELHOM, C. D. Effect of seed orientation, acid delinting, moisture level, and sample type on cottonseed fracture resistance. *Transactions of the ASABE*, v. 62, n. 4, p. 1045–1053, 2019.

ELMASRY, G. et al. Recent applications of multispectral imaging in seed phenotyping and quality monitoring—an overview. *Sensors*, v. 19, 1090, 2019.

ELMASRY, G. et al. Utilization of computer vision and multispectral imaging techniques for classification of cowpea (*Vigna unguiculata*) seeds. *Plant Methods*, v. 15, p. 24, 2019. DOI: 10.1186/s13007-019-0411-2.

GITELSON, A. A.; GRITZ, Y. e MERZLYAK, M. N. Relationships between leaf chlorophyll content and spectral reflectance. *Journal of Plant Physiology*, v. 160, p. 271–282, 2003.

HAN, J.; KAMBER, M. e PEI, J. *Data mining: concepts and techniques*. 3. ed. Waltham: Morgan Kaufmann, 2012.

HEMENDER, A. et al. Application of near infrared spectroscopy for seed quality evaluation: a review. *Agricultural Reviews*, v. 39, n. 3, p. 193–200, 2018.

HOURANT, P. et al. Oil and fat classification by selected bands of near-infrared spectroscopy. *Applied Spectroscopy*, v. 54, p. 1168–1174, 2000.

KOCZOŃ, P. et al. The analytical possibilities of FT-IR spectroscopy powered by vibrating molecules. *International Journal of Molecular Sciences*, v. 24, 1013, 2023.

- KRZYŻANOWSKI, F.; OHLSON, O. Série qualidade: plântulas anormais de algodão. *Seed News*, 2024.
- KUHN, M. Building predictive models in R using the caret package. *Journal of Statistical Software*, v. 28, n. 5, 2008.
- KUMAR, S. P. J.; PRASAD, S. R.; BANERJEE, R.; THAMMINENI, C. Seed birth to death: dual functions of reactive oxygen species in seed physiology. *Annals of Botany*, v. 116, n. 4, p. 663–668, 2015.
- KUSUMANINGRUM, D. et al. Non-destructive technique for determining the viability of soybean seeds using FT-NIR spectroscopy. *Journal of the Science of Food and Agriculture*, v. 98, p. 1734–1742, 2018.
- LARIOS, G. et al. Soybean seed vigor discrimination by using infrared spectroscopy and machine learning algorithms. *Analytical Methods*, v. 12, p. 4303–4309, 2020.
- LI, J. et al. Hyperspectral imaging for nondestructive evaluation of seed quality: a review. *Sensors*, v. 19, n. 18, p. 1–27, 2019. DOI: 10.3390/s19183911.
- LIMÃO, M. A. R. et al. Near-infrared spectroscopy and accelerated ageing in evaluating the vigour of lentil seeds. *Revista Ciência Agronômica*, v. 56, e202493547, 2025.
- LIU, W. et al. Rapid discrimination of high-quality watermelon seeds by multispectral imaging. *Journal of Applied Spectroscopy*, v. 85, p. 1044–1049, 2019.
- MARCOS-FILHO, J. *Fisiologia de sementes de plantas cultivadas*. 2. ed. Londrina: ABRATES, 2015.
- MARCOS-FILHO, J. Teste de envelhecimento acelerado. In: ABRATES (Ed.). *Vigor de sementes: conceitos e testes*. Londrina: ABRATES, 2020a. p. 182–244.
- MARCOS-FILHO, J. Teste de vigor: importância e utilização. In: ABRATES (Ed.). *Vigor de sementes: conceitos e testes*. Londrina: ABRATES, 2020b. p. 17–72.
- MASILAMANI, P. et al. Role of near-infrared spectroscopy in seed quality evaluation: a review. *Agricultural Reviews*, 2020.
- MATA, M. M. et al. Distinguishing cotton seed genotypes by vibrational spectroscopic methods. *Spectrochimica Acta Part A*, v. 266, 120399, 2022.

- MAYRINCK, L. G. et al. Use of near infrared spectroscopy in cotton seeds physiological quality evaluation. *Journal of Seed Science*, v. 42, 2020.
- MAYRINCK, L. G. et al. Use of near infrared spectroscopy in cotton seeds physiological quality evaluation. *Journal of Seed Science*, v. 42, p. 1–11, 2020. DOI: 10.1590/2317-1545v42227169.
- MEDEIROS, A. D. et al. Interactive machine learning for soybean seed and seedling quality classification. *Scientific Reports*, v. 10, 11267, 2020a.
- MEDEIROS, A. D. et al. Machine learning for seed quality classification. *Sensors*, v. 20, 4319, 2020b.
- MEDEIROS, A. D. et al. X-ray imaging and machine learning to evaluate the quality of crambe seeds. *Industrial Crops and Products*, v. 170, 113731, 2021.
- OLIVEIRA, G. R. et al. Spectroscopy technologies to screen peanut seeds with superior vigor through chemical fingerprinting. *Agronomy*, v. 14, 2529, 2024.
- OSBORNE, B. G.; FEARN, T.; HINDLE, P. H. *Practical NIR spectroscopy with applications in food and beverage analysis*. 2. ed. Harlow: Longman, 1993.
- OTSU, N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, v. 9, p. 62–66, 1979. DOI: 10.1109/TSMC.1979.4310076.
- OZAKI, Y.; MCCLURE, W. F.; CHRISTY, A. A. *Near-infrared spectroscopy in food science and technology*. New Jersey: Wiley, 2006.
- PASQUINI, C. Near infrared spectroscopy: a mature analytical technique with new perspectives. *Analytica Chimica Acta*, v. 1026, p. 8–36, 2018.
- PETRONILIO, A. C. P. et al. Smart and accurate: a new tool to identify stressed soybean seeds. *Smart Agricultural Technology*, v. 12, 101042, 2025.
- R CORE TEAM. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, 2025. Disponível em: <https://www.r-project.org/>. Acesso em: 5 nov. 2025.
- RIBEIRO, J. P. O. et al. FT-NIR and linear discriminant analysis to classify chickpea seeds. *Food Chemistry*, v. 342, 128324, 2021.

- RINNAN, Å.; BERG, F. van den; ENGELSEN, S. B. Review of preprocessing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry*, v. 28, p. 1201–1222, 2009.
- RUIZ-PEREZ, D. et al. So you think you can PLS-DA? *BMC Bioinformatics*, v. 21, p. 2, 2020.
- SANTANA, F. et al. Experimento didático de quimiometria para classificação de óleos vegetais. *Química Nova*, 2020.
- SCHÜTZ, D. et al. Fourier-transform near-infrared spectroscopy as a screening tool. *Food Control*, v. 136, 108892, 2022.
- SENA, M. M. e ALMEIDA, M. R. Quimiometria aplicada aos dados espectrais no infravermelho próximo. In: *Chemometrics applied to near-infrared spectral data*. Belo Horizonte: UFMG, 2018. p. 31–50.
- SHRESTHA, S.; DELEURAN, L. e GISLUM, R. Classification of tomato seed cultivars. *Journal of Spectral Imaging*, 2016.
- SILVA, C. et al. Autofluorescence-spectral imaging as an innovative method for rapid, non-destructive and reliable assessing of soybean seed quality. *Scientific Reports*, v. 11, 17834, 2021.
- SILVA, M. F. et al. Near infrared spectroscopy for the classification of vigor level of soybean seed. *Revista Ciência Agronômica*, v. 55, e20238703, 2024.
- SNIPES, C. E. e BASKIN, C. C. Influence of early defoliation on cotton yield. *Field Crops Research*, v. 37, p. 137–143, 1994.
- SOARES, J. M. et al. Classification of the physiological potential of soybean seed lots. *Journal of Seed Science*, v. 46, 2024.
- SOARES, J. M. et al. Low-cost system for multispectral image acquisition and its applicability to analysis of the physiological potential of soybean seeds. *Acta Scientiarum. Agronomy*, v. 45, 2023. DOI: 10.4025/actasciagron.v45i1.57060.
- SOARES, S. F. et al. Classification of individual cotton seeds with respect to variety using near-infrared hyperspectral imaging. *Analytical Methods*, v. 8, p. 8498–8505, 2016.
- SOUZA, L. R. et al. Near infrared spectroscopy and seedling image analysis. *Journal of Seed Science*, v. 45, 2023.

VRESAK, M. et al. Image-spectroscopy technology for seed testing. *PLoS One*, v. 11, e0152011, 2016. DOI: 10.1371/journal.pone.0152011.

WORKMAN J. R., J.; WEYER, L. *Practical guide to interpretive near-infrared spectroscopy*. Boca Raton: CRC Press, 2007.

XIA, Y. et al. Recent advances in emerging techniques for non-destructive detection of seed viability. *Artificial Intelligence in Agriculture*, v. 1, p. 35–47, 2019.

XU, R. et al. Use of near-infrared spectroscopy for soybean protein evaluation. *Spectrochimica Acta Part A*, v. 224, 117400, 2020.

CAPÍTULO III – Alterações bioquímicas, fisiológicas e análises espectrais em sementes de algodão submetidas à deterioração.

RESUMO

Sementes de algodão, por serem oleaginosas, são mais propensas à deterioração, que afeta negativamente a sua qualidade fisiológica. Considerando que a deterioração é um processo oxidativo que promove alterações nos constituintes químicos das sementes, devido principalmente à peroxidação lipídica, técnicas de espectroscopia e de análises multiespectrais podem ser promissoras para classificar lotes de sementes quanto à qualidade fisiológica, de acordo com o seu nível de deterioração. Objetivou-se verificar se é possível associar indicadores bioquímicos e análises espectrais na avaliação do potencial fisiológico de sementes de algodão deterioradas. Sementes de um lote da cv. 21065TLP foram envelhecidas artificialmente em BOD a 41° C, 100% UR por 0, 24, 48, 72, 96 e 120 horas a 41°C. Após cada período, foram submetidas aos seguintes testes de avaliação da qualidade fisiológica: germinação e primeira contagem de germinação, condutividade elétrica e crescimento de plântulas. Foram também realizadas análises bioquímicas: enzimas do sistema antioxidante (catalase - CAT, superóxido dismutase - SOD, e peroxidase - POX), quantificação de peróxido de hidrogênio e acúmulo de malonaldeído (MDA). O delineamento utilizado foi o inteiramente casualizado, com quatro repetições. As médias obtidas para cada tratamento de envelhecimento artificial foram submetidas à análise de regressão. Para a aquisição dos espectros NIR, 200 sementes de cada tratamento foram analisadas individualmente em espectômetro FT-NIR, na faixa de comprimento de onda de 1.000 a 2.500 nm. Imagens multiespectrais também foram obtidas nas bandas espectrais de 395, 460, 520, 585, 620, 740, 850 e 940 nm em equipamento protótipo composto por câmera digital e lâmpadas LED. Em seguida, estas mesmas sementes foram avaliadas quanto à germinação. Foram utilizados os algoritmos de pré-processamento: *Standard Normal Variate* (SNV); *Multiplicative Scatter Correction* (MSC); e as 1ª e 2ª derivadas de Savitzky-Golay (SG) para os dados obtidos no FT-NIR e para os dados multiespectrais o SNV e o MSC. Três algoritmos de classificação foram aplicados: *Partial Least Squares Discriminant Analysis* (PLS-DA), *Random Forest* (RF) e *Neural Networks* (NN). Os tratamentos foram classificados em três categorias com base nos resultados do teste de germinação: alto (germinação $\geq 90\%$), médio (germinação entre 80–89%) e baixo potencial fisiológico (germinação $\leq 79\%$). A atividade das enzimas do sistema antioxidantes SOD, CAT e POX reduziu a medida em que ocorria o aumento do tempo de envelhecimento artificial, sendo

inversamente proporcional ao acúmulo de H₂O₂ e MDA. A espectroscopia no infravermelho próximo é eficiente para classificar sementes de algodão, com destaque para o algoritmo PLS-DA e o pré-processamento com a 2ª derivada Savitzky-Golay, alcançando acurácia de 100%. O modelo *Random Forest* e pré-processamento MSC foi eficiente para classificar sementes de algodão pela técnica de análise de imagens multiespectrais, alcançando acurácia de até 92%.

Palavras-chave: deterioração de sementes, bioquímica, FT-NIR, imagens multiespectrais, *Gossypium hirsutum* L.

ABSTRACT

Cotton seeds, because they are oilseeds, are more prone to deterioration, which negatively affects their physiological quality. Considering that deterioration is an oxidative process that promotes changes in seed chemical constituents, mainly due to lipid peroxidation, spectroscopy and multispectral analysis techniques may be promising for classifying seed lots according to their physiological quality based on their level of deterioration. This study aimed to verify whether it is possible to associate biochemical indicators and spectral analyses in the evaluation of the physiological potential of deteriorated cotton seeds. Seeds from one lot of the cultivar 21065TLP were artificially aged in a BOD chamber at 41 °C and 100% relative humidity for 0, 24, 48, 72, 96, and 120 hours. After each period, seeds were subjected to the following physiological quality tests: germination and first germination count, electrical conductivity, and seedling growth. Biochemical analyses were also performed, including antioxidant enzyme activity (catalase – CAT, superoxide dismutase – SOD, and peroxidase – POX), as well as quantification of hydrogen peroxide (H₂O₂) and malondialdehyde (MDA) accumulation. A completely randomized design with four replications was used. The means obtained for each artificial aging treatment were subjected to regression analysis. For NIR spectral acquisition, 200 seeds per treatment were individually analyzed using an FT-NIR spectrometer in the wavelength range of 1,000 to 2,500 nm. Multispectral images were also obtained at spectral bands of 395, 460, 520, 585, 620, 740, 850, and 940 nm using a prototype equipped with a digital camera and LED lamps. These same seeds were subsequently evaluated for germination. The following preprocessing algorithms were applied to FT-NIR data: Standard Normal Variate (SNV), Multiplicative Scatter Correction (MSC), and first- and second-order Savitzky–Golay derivatives. For multispectral data, SNV and MSC were used. Three classification algorithms were applied: Partial Least Squares Discriminant Analysis (PLS-DA), Random Forest (RF), and Neural Networks (NN). Treatments were classified into three categories based on germination results: high (germination \geq 90%), medium (germination between 80–89%), and low physiological potential (germination \leq 79%). The activity of antioxidant enzymes SOD, CAT, and POX decreased as the artificial aging period increased, showing an inverse relationship with the accumulation of H₂O₂ and MDA. Near-infrared spectroscopy proved to be efficient for classifying cotton seeds, with the PLS-DA algorithm combined with second-order Savitzky–Golay preprocessing showing the best performance, achieving accuracy of up to 100%. The Random Forest model with MSC preprocessing was efficient for classifying cotton seeds using multispectral image analysis, reaching accuracy of up to 92%.

Keywords: seed deterioration, chemical composition, FT-NIR, multispectral images, *Gossypium hirsutum* L.

INTRODUÇÃO

O cultivo de algodão no Brasil destaca-se pelo uso de tecnologias modernas, que requerem altos investimentos (ABRAPA, 2025). Neste contexto, as exigências com relação à alta qualidade e desempenho em campo das sementes comercializadas são cada vez maiores, devido ao seu alto valor comercial. O elevado teor de óleo das sementes de algodão contribui para uma maior propensão ao processo de deterioração, o que demanda cuidados específicos durante as etapas de produção e pós-colheita, para preservação da sua qualidade fisiológica (Shahein et al., 2022).

Bioquimicamente, o processo de deterioração é provocado pela produção excessiva de espécies reativas de oxigênio (EROs) decorrentes do estresse oxidativo que acarretam uma série de alterações metabólicas como a peroxidação de lipídeos, oxidação de proteínas, danos a ácidos nucleicos, inibição enzimática e morte celular (Pinheiro et al., 2023; Mittler, 2002; Krzyzanowski; Dias; França-Neto, 2022). As EROs são subprodutos naturais do metabolismo formadas a partir da redução parcial do oxigênio, destacando-se o radical superóxido ($O_2^{\bullet-}$), peróxido de hidrogênio (H_2O_2), radical hidroxila (OH^{\bullet}), dentre outros (Mittler, 2002; Parida e Das, 2005; Ray; Huang; Tsuji, 2012). O estresse oxidativo ocorre quando há um desequilíbrio entre a geração de EROs e a eficiência dos sistemas antioxidantes responsáveis por removê-las ou neutralizá-las (Cembrowska-Lech e Rybak, 2023; Mittler, 2017). Nesse contexto, o sistema antioxidante enzimático, constituído pelas enzimas superóxido dismutase (SOD), catalase (CAT), peroxidase do ascorbato (APX) e peroxidases (POX), dentre outras, desempenha papel essencial na manutenção da homeostase celular (Bailey, 2023; Sachdev et al., 2021).

Em sementes de soja, a deterioração na fase de pré-colheita reduziu a atividade das enzimas antioxidativas CAT, APX e POX bem como a germinação e o vigor, observando-se também aumento no acúmulo de H_2O_2 e na peroxidação de lipídeos determinada por meio do conteúdo de malonaldeído (Pinheiro et al., 2023). Semelhantemente, Wang et al. (2012) observaram aumento na produção de EROs e peroxidação de lipídeos em sementes sob condições de alta temperatura e umidade relativa. Estudos com sementes algodão envelhecidas artificialmente, evidenciaram que a redução da germinação apresentou alta correlação com o aumento do conteúdo de H_2O_2 e peroxidação lipídica, bem como com a redução das atividades das enzimas SOD, CAT e POX (Goel et al., 2002). A deterioração das sementes de algodão durante o envelhecimento está diretamente relacionada à redução da atividade de enzimas antioxidantes que têm função de reduzir os efeitos deletérios ocasionados pelo excesso de EROs e peroxidação lipídica (Zheng et al., 2024). Assim, a redução da atividade destas enzimas pode

intensificar o estresse oxidativo, acelerando o processo de deterioração e o declínio do vigor das sementes, conforme observado por Xing et al. (2023).

Considerando que a deterioração é um processo oxidativo que promove alterações nos constituintes químicos das sementes, devido principalmente à peroxidação lipídica, técnicas de espectroscopia e de análises multiespectrais podem ser promissoras para classificar lotes de sementes quanto à qualidade fisiológica, de acordo com o seu nível e deterioração (Xia et al., 2019), principalmente quando são associadas à inteligência artificial (Medeiros et al., 2020a; Petronilio et al., 2025). Tais tecnologias têm como vantagem a rapidez e automação, permitindo o processamento de um grande número de lotes, são não destrutivas, não requer preparo prévio da amostra e são objetivas.

O FT-NIR tem como princípio a interação da radiação eletromagnética na faixa de 780 a 2500 nm com os grupos funcionais das moléculas presentes nas sementes (O–H, C–H e N–H), apresentando relação direta com a composição química e estrutural das sementes como água, lipídeos, carboidratos e proteínas (Xia et al., 2019). Os recentes avanços da utilização dessa técnica confirmam que é possível aplicá-la para a classificação de lotes de sementes quanto ao potencial fisiológico, como constatado em soja (Soares et al., 2024), milho (Schutz et al., 2022), grão-de-bico (Ribeiro et al., 2021), braquiária (Souza et al., 2023), trigo (Xia et al., 2019) e melancia (Yasmin et al., 2019). Em algodão, a técnica se mostrou eficiente para distinguir diferentes genótipos (Mata et al., 2022), estimar o potencial fisiológico de sementes com e sem línter (Mayrinck et al., 2020) e avaliar a composição química das sementes (Zhuang et al., 2023).

Em relação as imagens multiespectrais, essa tecnologia possibilita a avaliação detalhada de características físicas e químicas das sementes, fornecendo dados espaciais e espectrais utilizando comprimentos de onda específicos que permitem avaliar atributos de qualidade como estrutura de superfície, textura, composição química e, potencialmente, aspectos fisiológicos (Medeiros et al., 2020a). Assim, a luz, ao interagir com a matéria, é absorvida, refletida ou transmitida em comprimentos de onda específicos, dependendo da composição química da amostra (Galletti et al., 2020; Petronilio et al., 2025). Em sementes, as imagens multiespectrais criam uma "assinatura química", permitindo extrair marcadores ópticos referentes à composição química, conteúdo de fluoróforos como clorofila *a*, clorofila *b*, lignina e atributos físicos das sementes como cor, brilho, textura e área (França-Silva et al., 2023). Essas informações espectrais estão intrinsecamente ligadas ao potencial fisiológico das sementes, permitindo uma avaliação não destrutiva, precisa e rápida (Petronilio et al., 2025). O uso dessa

técnica foi comprovado para a identificação de sementes de diferentes cultivares de tomate (Shrestha et al., 2015) e trigo (Vresak et al., 2016), na classificação de diferentes níveis de vigor em sementes de soja utilizando autofluorescência (Barboza da Silva et al., 2021), na avaliação do potencial fisiológico de sementes de cenoura e tomate via reflectância e fluorescência de clorofila (Galletti et al., 2020) e na identificação de sementes de soja produzidas sob condições favoráveis e de estresse, por meio da fluorescência de clorofila (Petronilio et al., 2025).

Sementes deterioradas apresentam alterações bioquímicas e estruturais que modulam a absorção e a reflectância da luz nos diferentes comprimentos de onda, tornando possível separar sementes de alto e baixo vigor (França-Silva et al., 2023; Petronilio et al., 2025). Assim, tanto o FT-NIR como a análise de imagens multiespectrais podem atuar como uma alternativa rápida e não destrutiva para avaliar o nível de deterioração e associar ao potencial fisiológico das sementes, principalmente quando combinados com modelos de aprendizado de máquina.

Diante do exposto, objetivou-se avaliar as alterações bioquímicas e fisiológicas e análises espectrais em sementes de algodão submetidas à deterioração.

MATERIAL E MÉTODOS

A pesquisa foi conduzida no Laboratório de Pesquisa de Sementes, no Laboratório de Agroenergia do Departamento de Agronomia e no Laboratório de Crescimento e Desenvolvimento de Plantas da Universidade Federal de Viçosa, Viçosa, MG, Brasil. Foram utilizadas sementes comerciais de algodão, cultivar 21065TLP.

Inicialmente, as sementes foram submetidas a determinação do grau de umidade, determinado em estufa a 105 °C por 24 h (Brasil, 2025), sendo os resultados expressos em porcentagem.

Para se obter sementes com diferentes níveis de deterioração, estas foram envelhecidas artificialmente por diferentes períodos, conforme metodologia descrita para o teste de envelhecimento acelerado (Marcos-Filho, 2020). Para cada período, foram utilizadas 250 sementes, distribuídas em camada uniforme sobre tela metálica acoplada em caixa plástica tipo gerbox, contendo ao fundo 40 mL de água destilada. As caixas foram devidamente tampadas para se obter aproximadamente 100% UR em seu interior e, em seguida, mantidas em câmara tipo BOD a temperatura de 41 °C por períodos de 24, 48, 72, 96 e 120 horas. Para o tratamento controle, sementes não envelhecidas foram utilizadas. O grau de umidade inicial das sementes foi de 9,5. Após o envelhecimento artificial, o grau de umidade das sementes atingiu 31,6% em 24 horas; 33,6% em 48 horas; 35,5% em 72 horas; 38,3% em 96 horas; e 40,8% em 120 horas.

As sementes de cada tratamento foram então secas em ambiente de laboratório até atingirem grau de umidade em torno de 9%, para que fossem realizadas os testes fisiológicos, bioquímicos e análises de espectroscopia e de imagens multiespectrais.

Para as análises de espectroscopia no infravermelho próximo (FT-NIR) e imageamento multiespectrais, as sementes de cada tratamento foram mapeadas e identificadas, para posterior condução do teste de germinação.

Caracterização do Potencial Fisiológico:

Germinação (G) e primeira contagem de germinação (PCG): quatro repetições de 50 sementes foram distribuídas em papel toalha umedecido com água na quantidade de 2,5 vezes o peso do papel seco. Os rolos foram confeccionados e mantidos em germinador a 20°C. Para a primeira contagem de germinação calculou-se a porcentagem de plântulas normais obtidas no quarto dia após a sementeira, e a porcentagem de germinação foi determinada aos 12 dias após a sementeira (Brasil, 2009).

Condutividade elétrica: quatro repetições de 50 sementes foram pesadas (precisão de 0,001 g) e colocadas em copos plásticos contendo o volume de 70 mL de água destilada. Os copos foram mantidos em incubadora BOD a 25 °C por 24 h. Após esse período, foi realizada a leitura da condutividade elétrica por meio de condutivímetro (Digimed CD 21) e os resultados foram expressos em $\mu\text{S}\cdot\text{cm}^{-1}\cdot\text{g}^{-1}$ (Vieira e Marcos-Filho, 2020).

Crescimento de plântulas: realizado com 8 repetições de 20 sementes por tratamento. As sementes foram distribuídas em uma linha traçada no terço superior do papel de germinação no sentido longitudinal. As sementes foram posicionadas de forma que o hilo ficasse direcionado para a parte inferior do papel e os rolos foram acondicionados em sacos plásticos, colocados verticalmente no germinador por 4 dias a 25 °C (Krzyzanowski et al., 2020). Ao final do período, as plântulas foram transferidas do papel de germinação para a bandeja de acrílico transparente acopladas no equipamento do GroundEye, de forma que não se tocassem. A captura das imagens do teste de crescimento foi realizada por uma câmera profissional de alta resolução acoplada ao equipamento GroundEye. O processamento das imagens foi realizado por meio do *software* GroundEye obtendo-se valores de comprimento do hipocótilo, comprimento da raiz e comprimento total das plântulas de algodão.

Análises bioquímicas

Preparo das amostras

Inicialmente foi realizada a embebição das sementes dos diferentes lotes de algodão para remoção do tegumento e posteriormente a realização das determinações enzimáticas. Para tanto,

foram utilizadas 4 repetições com 50 sementes de cada lote, colocadas para embeber em rolo de papel toalha umedecido com água destilada em quantidade equivalente a 2,5 vezes o peso do papel seco e mantidos em germinador a 25 °C no período de 24 h. Após esse período, o tegumento das sementes foi removido.

Posteriormente, o material vegetal foi congelado em nitrogênio líquido e liofilizado. Em seguida, o material liofilizado foi macerado em moinho de bolas e armazenado em tubos do tipo eppendorf, mantidos dentro de um dessecador para não ocorrer perda de umidade até o momento de quantificação da atividade enzimática.

Os teores de proteínas dos extratos enzimáticos foram determinados pelo método de Bradford (1976), utilizando BSA como padrão. 2,5 µL do extrato enzimático, que foi adicionado a 1 mL do reagente de Bradford, seguido de agitação. Após 20 minutos, foi realizada a leitura da absorbância da amostra em espectrofotômetro a 595 nm.

Enzimas do sistema antioxidante: Para a determinação da atividade das enzimas catalase (CAT, EC 1.11.1.6), superóxido dismutase (SOD, EC 1.15.1.1) e peroxidases (POX, EC 1.11.1.7), 15 mg de amostra liofilizadas foram extraídas por homogeneização de 1 mL do meio de extração (tampão de fosfato de potássio 0,1 M e pH 6,8; contendo ácido etilenodiaminotetracético (EDTA) 0,1 mM, fluoreto de fenilmetilsulfonil (PMSF) 1 mM e polivinilpirrolidona (PVPP) 1% (p/v) (Peixoto et al., 1999). A mistura foi centrifugada a 12.000 xg durante 15 min, a 4 °C e o sobrenadante foi utilizado como extrato enzimático.

A atividade da SOD foi conduzida a 25 °C, em câmara de reação sob iluminação durante por 5 min e medida a 560 nm. Uma unidade de SOD foi definida como a quantidade de enzima necessária para inibir em 50% a fotoredução do NBT e os resultados foram expressos em U min⁻¹ mg⁻¹ proteína (Beauchamp e Fridovich, 1971).

Para CAT, a atividade foi determinada medindo-se o decréscimo na absorbância, no primeiro minuto de reação a 240 nm, a 30 °C e esta calculada utilizando-se o coeficiente de extinção molar de 36 M⁻¹ cm⁻¹ (Anderson et al., 1995) e os resultados expressos em µmol de H₂O₂ min⁻¹ mg⁻¹ proteína.

Já a atividade da POX foi medida o incremento da absorbância durante o primeiro minuto de reação a 420 nm, a 25°C, sendo a atividade calculada utilizando-se o coeficiente de extinção molar de 2,47 mM⁻¹ cm⁻¹ (Chance e Maehley, 1955) e os resultados expressos em µmol de purpurogalina min⁻¹ mg⁻¹ proteína.

Quantificação de peróxido de hidrogênio: Foram utilizadas amostras contendo 0,15 g de embriões (cotilédones e eixo embrionário) trituradas em nitrogênio líquido e

homogeneizadas em 2,0 mL de tampão fosfato de potássio 50 mM, pH 6,5, contendo hidroxilamina 1 mM, seguido de centrifugação a 14.000 g, por 15 min, 4°C, e coleta do sobrenadante (Kuo; Kao, 2003). Alíquotas de 100 µL do sobrenadante foram adicionadas a um meio de reação constituído de sulfato ferroso amoniacal 250 µM, ácido sulfúrico 25 mM, laranja de xilenol 250 µM e sorbitol 100 mM, em volume final de 2 mL (Gay; Gebicki, 2000) homogeneizadas e mantidas no escuro por 30 min. A determinação da absorvância a 560 nm e a quantificação de H₂O₂ foi realizada com base em curva de calibração, com a utilização de concentrações de peróxido como padrão. Brancos para os reagentes e os extratos vegetais foram preparados em paralelo e subtraídos da amostra.

Peroxidação de lipídeos: A peroxidação de lipídeos foi quantificada pelo acúmulo de malonaldeído (MDA), conforme descrito por Cakmak e Horst (1991). Utilizou-se 0,15 g de embriões (cotilédones e eixo embrionário) que foram macerados em ácido tricloroacético (TCA 0,1%, p/v). Depois da centrifugação (19.000 g, 15 min., 4 °C), 500 µL do sobrenadante foi coletado e adicionado a 1,5 mL de solução de ácido tiobarbitúrico (TBA 0,5% em TCA 20%). No branco, em vez da amostra, foi adicionado 500 µL de TCA 0,1%. As amostras e o branco foram incubados por 30 min a 90 °C em banho-maria, com agitação. Após 30 minutos, a reação foi parada em gelo e realizou-se outra centrifugação (19.000 g, 15 min., 4 °C). As leituras foram feitas em espectrofotômetro a 532 nm. A concentração de MDA foi calculada utilizando-se o coeficiente de extinção molar de 155 mM⁻¹ cm⁻¹ e os resultados expressos em nmol.g⁻¹ de matéria fresca.

Espectroscopia no Infravermelho Próximo e Análise de Imagens Multiespectrais

Aquisição do espectro FT-NIR: Para cada período de EA, 200 sementes foram selecionadas aleatoriamente, totalizando 1.200 sementes avaliadas individualmente. Os dados espectrais de cada semente foram obtidos utilizando um espectrômetro de Infravermelho Próximo com Transformada de Fourier (FT-NIR) (Thermo Scientific Antaris II). Os espectros foram expressos em absorbância (log 1/R). Cada espectro de semente foi composto por 3.111 pontos dentro da faixa de comprimento de onda de 1.000 a 2.500 nm. Foram realizadas trinta varreduras sucessivas, e o espectro médio de cada semente foi calculado.

Análise de imagens multiespectrais: As mesmas sementes utilizadas para a análise FT-NIR foram submetidas à aquisição de imagens multiespectrais. Foi utilizado um protótipo equipado com uma câmera digital e LEDs (Viçosa, Minas Gerais, Brasil) para a captura das imagens, fornecendo imagens UV-VIS-NIR nos comprimentos de onda de 395, 460, 520, 585,

620, 740, 850 e 940 nm, totalizando oito bandas espectrais. As imagens foram capturadas e salvas no formato TIFF (Brochado et al., 2024).

Para a análise das imagens e extração dos valores de nível de cinza, foi utilizado o software ImageJ (versão ImageJ2, Bethesda, Maryland, EUA). As imagens foram processadas para selecionar a região de interesse utilizando o método de segmentação *Threshold* com Otsu 1979 e os valores médios de tons de cinza dos pixels individuais das sementes foram obtidos para cada banda espectral. Após a obtenção das curvas espectrais com base nos valores médios de cinza, as curvas foram plotadas e o processamento dos dados foi realizado para desenvolver modelos de classificação utilizando o pacote Caret (Kuhn, 2008).

Após a aquisição dos espectros FT-NIR e das imagens multiespectrais, as sementes de cada tratamento foram submetidas ao teste de germinação.

Delineamento Experimental e Análise Estatística

Caracterização do Potencial Fisiológico e Análises Bioquímicas: Para os testes de caracterização do potencial fisiológico, enzimas do sistema antioxidante, conteúdo de peróxido de hidrogênio e de malonaldeído, utilizou-se o delineamento inteiramente casualizado, com seis tratamentos de períodos de envelhecimento artificial (controle, 24, 48, 72, 96 e 120 h) e quatro repetições. Os dados foram testados quanto à distribuição normal dos erros pelo teste de Shapiro-Wilk e à homogeneidade de variâncias pelo teste de Bartlett. As médias dos tratamentos foram submetidas à análise de regressão, ajustando-se modelos polinomiais em função dos períodos de envelhecimento artificial. A escolha do modelo foi baseada na significância dos coeficientes pelo teste t ($p < 0,05$), no coeficiente de determinação (R^2) e na adequação biológica do modelo. Foi realizada também a análise de correlação de Pearson (r) para as variáveis de qualidade fisiológica e análises bioquímicas. As análises estatísticas foram realizadas utilizando o software R, versão 4.3.3 (R Core Team, 2025).

Espectroscopia no Infravermelho Próximo e Análise de Imagens Multiespectrais:

Algoritmos de Pré-processamento: Os dados espectrais originais de NIR foram pré-processados utilizando os seguintes métodos: *Standard Normal Variate* (SNV); *Multiplicative Scatter Correction* (MSC); e as 1ª e 2ª derivadas de Savitzky-Golay (SG) com janela de 7 pontos. Para os dados extraídos das imagens multiespectrais (valores de cinza), foi aplicado apenas os pré-processamentos *Standard Normal Variate* (SNV) e *Multiplicative Scatter Correction* (MSC)

Algoritmos de Classificação: Três modelos de classificação foram aplicados aos espectros obtidos por diferentes técnicas de pré-processamento e aos valores de cinza extraídos

das imagens multiespectrais: *Partial Least Squares Discriminant Analysis* (PLS-DA), *Random Forest* (RF) e *Neural Networks* (NN). Para os algoritmos PLS-DA, Random Forest e Rede Neural, foi utilizado o pacote Caret (Kun, 2008). O algoritmo *Random Forest* foi executado com 200 árvores de decisão, enquanto a Rede Neural utilizou 20 neurônios na camada oculta.

Desenvolvimento dos Modelos de Classificação: Para o FT-NIR, foram estabelecidas três classes com base nos resultados do teste de germinação com as sementes submetidas a diferentes níveis de deterioração pelo envelhecimento artificial: alto potencial fisiológico (germinação $\geq 90\%$), médio potencial fisiológico (germinação entre 80–89%) e baixo potencial fisiológico (germinação $\leq 79\%$). Assim, cada classe de potencial fisiológico compreendeu 400 espectros. Para a análise multiespectrais, o banco de dados utilizado foi de 100 sementes por período de envelhecimento, totalizando assim 200 espectros de escala de cinza para cada uma das três classes.

Validação dos Modelos: Os modelos foram desenvolvidos utilizando 70% dos dados para treinamento e 30% para validação. Os modelos PLS-DA, Random Forest e Rede Neural foram avaliados com base nas métricas de acurácia (Eq. 1) e kappa (κ) (Eq. 2), conforme as equações abaixo:

$$\text{Acurácia} = \frac{\text{corretos}}{\text{total}} \quad (1)$$

$$K = \frac{P_o - P_e}{1 - P_e} \quad (2)$$

onde P_o é a proporção de concordância observada e P_e é a proporção de concordância esperada.

RESULTADOS E DISCUSSÃO

Caracterização do potencial fisiológico

Observa-se, na Figura 1A, a redução do poder germinativo das sementes e da velocidade de germinação (primeira contagem de germinação) à medida que aumenta o período de envelhecimento artificial. Comentar sobre o ajuste dos modelos (Figura 1A) Comportamento semelhante foi observado para a primeira contagem de germinação (Figura 1A). Pelo teste de condutividade elétrica (Figura 1B), observou-se aumento dos valores com o aumento do período de envelhecimento acelerado, indicando redução no vigor, o que pode ser atribuído ao aumento na desestruturação das membranas celulares e consequente lixiviação de solutos durante a embebição (Vieira e Marcos-Filho, 2020).

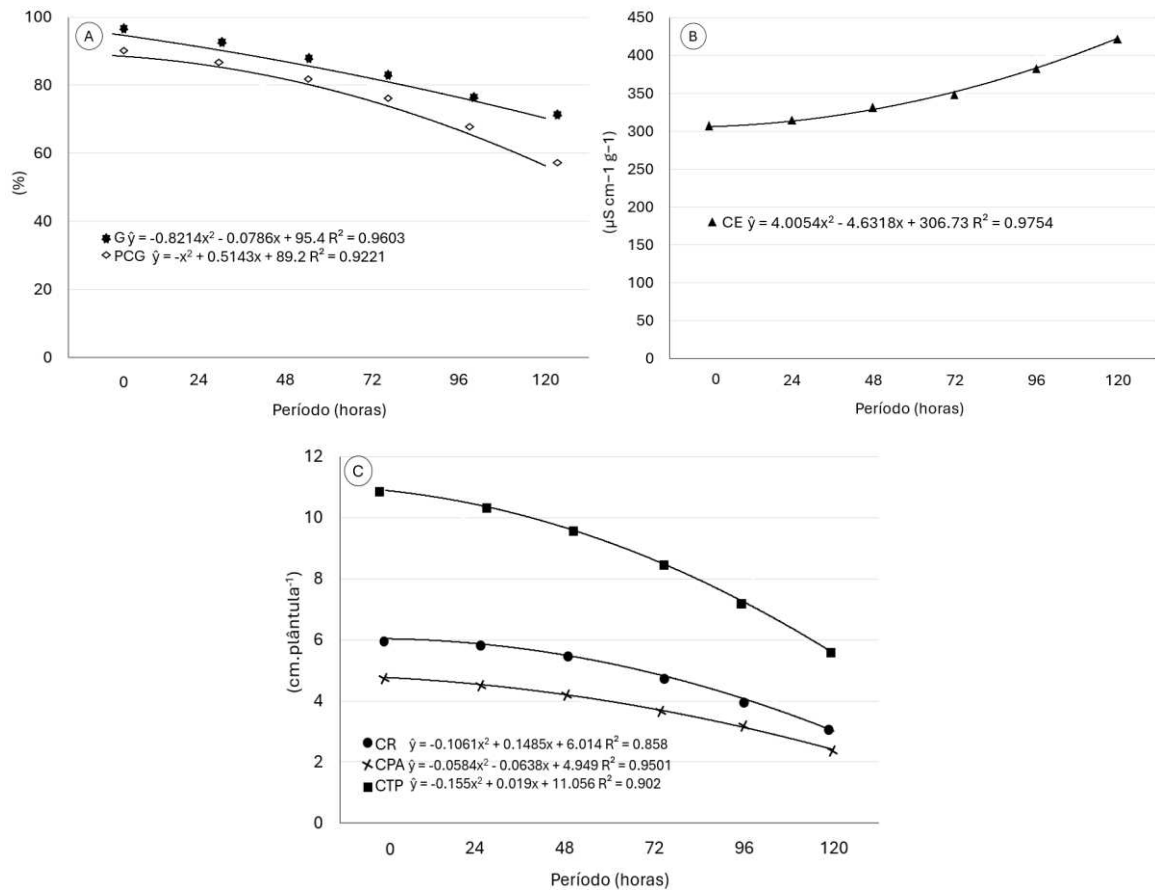


Figura 1. Germinação (G) e primeira contagem de germinação (PCG) – (A); condutividade elétrica (CE) – (B); comprimento da parte aérea (CPA); comprimento radicular (CR) e comprimento total de plântulas (CTP) – (C) de sementes de algodão após envelhecimento artificial.

Para as variáveis de comprimento da parte aérea (CPA), comprimento radicular (CR) e comprimento total de plântulas (CTP) (Figura 1C) observa-se redução do crescimento das plântulas a medida em que ocorreu aumento do período de envelhecimento artificial. A redução foi mais acentuada no CTP, principalmente a partir de 24 h de envelhecimento, enquanto para CPA e CR, a redução foi mais evidente a partir do período de 48 h.

De modo geral, os resultados obtidos para variáveis fisiológicas demonstram que a deterioração provocada pelo envelhecimento artificial reduziu o potencial fisiológico das sementes. Observa-se que essa redução foi proporcional ao aumento do tempo de exposição às condições de alta temperatura e umidade relativa do envelhecimento artificial, sendo que o maior período (120 horas) resultou em redução mais acentuada na germinação (Figura 1A), maior desestruturação das membranas e liberação de exsudados (Figura 1B), como também reduções drásticas no crescimento das plântulas (Figura 1C).

A deterioração e o vigor de sementes são processos fisiológicos inversamente relacionados, pois à medida que ocorre o aumento da deterioração há decréscimo do vigor (Krzyzanowski; Dias; França-Neto, 2022). Sementes envelhecidas, seja naturalmente ou artificialmente, passam por uma série de alterações que resultam na perda do vigor e da germinação. Delouche e Baskin (1973) relataram que um dos primeiros eventos que caracterizam o processo de deterioração é a desestruturação das membranas celulares, uma vez que sementes deterioradas liberam maior quantidade de exsudados, conforme observado pelo teste de condutividade elétrica (Figura 1B). Posteriormente, a atividade de biossíntese de moléculas essenciais no metabolismo celular é comprometida, ocasionando assim a redução na velocidade de germinação, como observado pelos resultados de primeira contagem de germinação (Figura 1A) com o aumento do período de envelhecimento artificial (Limão et al., 2025; Andrade et al., 2020; Wang et al., 2020; Yasmin et al., 2019). Em seguida, ocorre a redução do crescimento das plântulas (Figura 1C), que também é um reflexo da redução da biossíntese das moléculas essenciais do metabolismo celular. Observa-se que houve redução de 6,5 cm no comprimento total de plântulas ao se comparar o controle (sem envelhecimento) com sementes envelhecidas por 120 horas (Figura 1C).

Houve redução linear da atividade das enzimas do sistema antioxidante SOD (Figura 2A), CAT (Figura 2B) e POX (Figura 2C) com o aumento do período de envelhecimento artificial das sementes. Estes resultados reforçam a importância da atividade antioxidante destas enzimas protegendo as células contra o dano oxidativo causado pela deterioração (Mittler, 2017). A redução da atividade destas enzimas com o avanço da deterioração provavelmente contribuiu para o decréscimo da germinação (Figura 1A) e do vigor (Figuras 1B e 1C) das sementes. Os resultados da análise de correlação (Figura 3) reforçam esta afirmativa, onde é possível observar altos valores positivos para os coeficientes de correlação entre as variáveis de qualidade fisiológica e a atividade de enzimas antioxidantes (Figura 3).

A Superóxido Dismutase (SOD) é considerada uma das primeiras linhas de defesa celular contra o estresse oxidativo (Gill et al., 2015). Sua principal função é catalisar a dismutação do radical superóxido (O_2^-) convertendo-o em H_2O_2 e O_2 que são moléculas menos reativas (Pinheiro et al., 2023; Bailey, 2004). Neste contexto, o decréscimo na atividade da SOD tem relação direta com a deterioração de sementes, uma vez que o acúmulo de O_2^- é associado com a redução do potencial fisiológico de sementes (Bailey, 2004; Pinheiro et al., 2023).

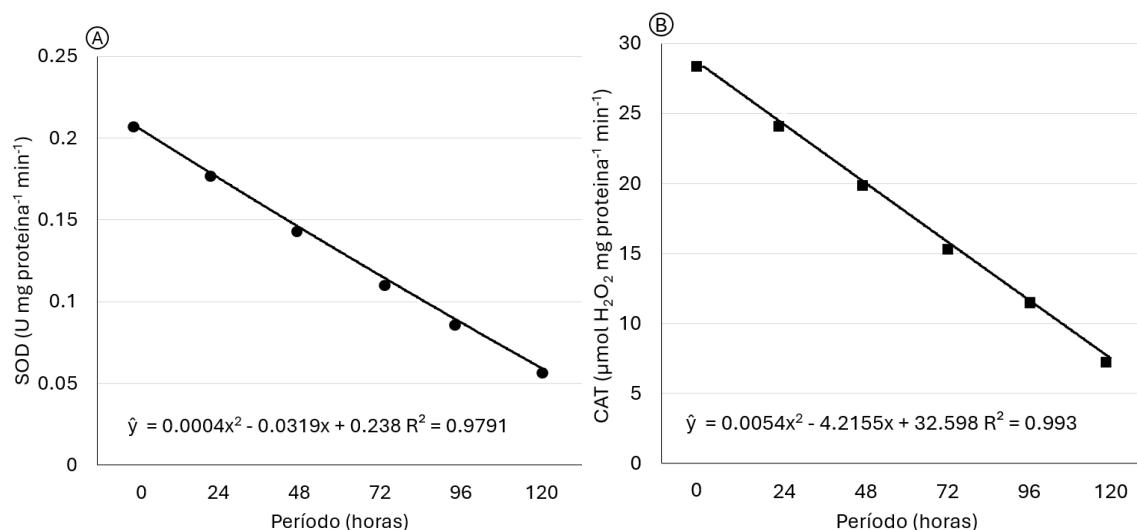
Já a CAT tem como principal função a conversão do H_2O_2 em H_2O e O_2 , prevenindo assim a formação de radicais hidroxila altamente reativos (Kibinza et al., 2011). Essa reação é

essencial para a manutenção do equilíbrio redox celular e para a integridade das estruturas biológicas. Em sementes, a CAT atua nos processos metabólicos e sua atividade está associada à tolerância das células a diferentes condições de estresse (Gill e Tuteja, 2010). Considerando que o envelhecimento artificial por altas temperaturas e alta umidade relativa acarretou a deterioração e, conseqüentemente, queda no vigor das sementes, espera-se o declínio da atividade da CAT (Bailey, 2004), conforme observado na Figura 1B.

A POX catalisa a decomposição do H_2O_2 juntamente com a SOD e a CAT, reduzindo o acúmulo de radicais livres, atuando na manutenção da homeostase redox. Em sementes deterioradas, a redução da atividade dessas enzimas favorece o acúmulo de ROS, intensificando a peroxidação lipídica, comprometendo a integridade das membranas celulares e resultando na redução do vigor e da germinação (Zhou et al., 2019).

Pode-se afirmar, portanto, que com o aumento do período de envelhecimento acelerado e conseqüente deterioração, o sistema de enzimas antioxidantes perde eficiência. Com isto, há um aumento na produção de espécies reativas de oxigênio (EROs), o que pode causar danos a lipídios, proteínas e ácidos nucleicos (Bailey, 2004).

Lotes de maior vigor, que apresentam reservas intactas, conseqüentemente maior germinação apresentaram maior atividade da CAT, SOD e APX em sementes de lentilha (Limão et al., 2024). Em soja, a atividade das enzimas antioxidantes como SOD, CAT e POX foi afetada conforme o nível de potencial fisiológico sob deterioração por intemperismo na fase de pré-colheita (Pinheiro et al., 2023). Outros estudos também já comprovaram a relação das enzimas do sistema antioxidativo CAT, SOD e POX com a deterioração e a redução do potencial fisiológico de sementes de grão-de-bico (Araújo et al., 2021), feijão mungo-verde (Sharma et al., 2018), milho (Borba et al., 2014) e soja (Pinheiro et al., 2023).



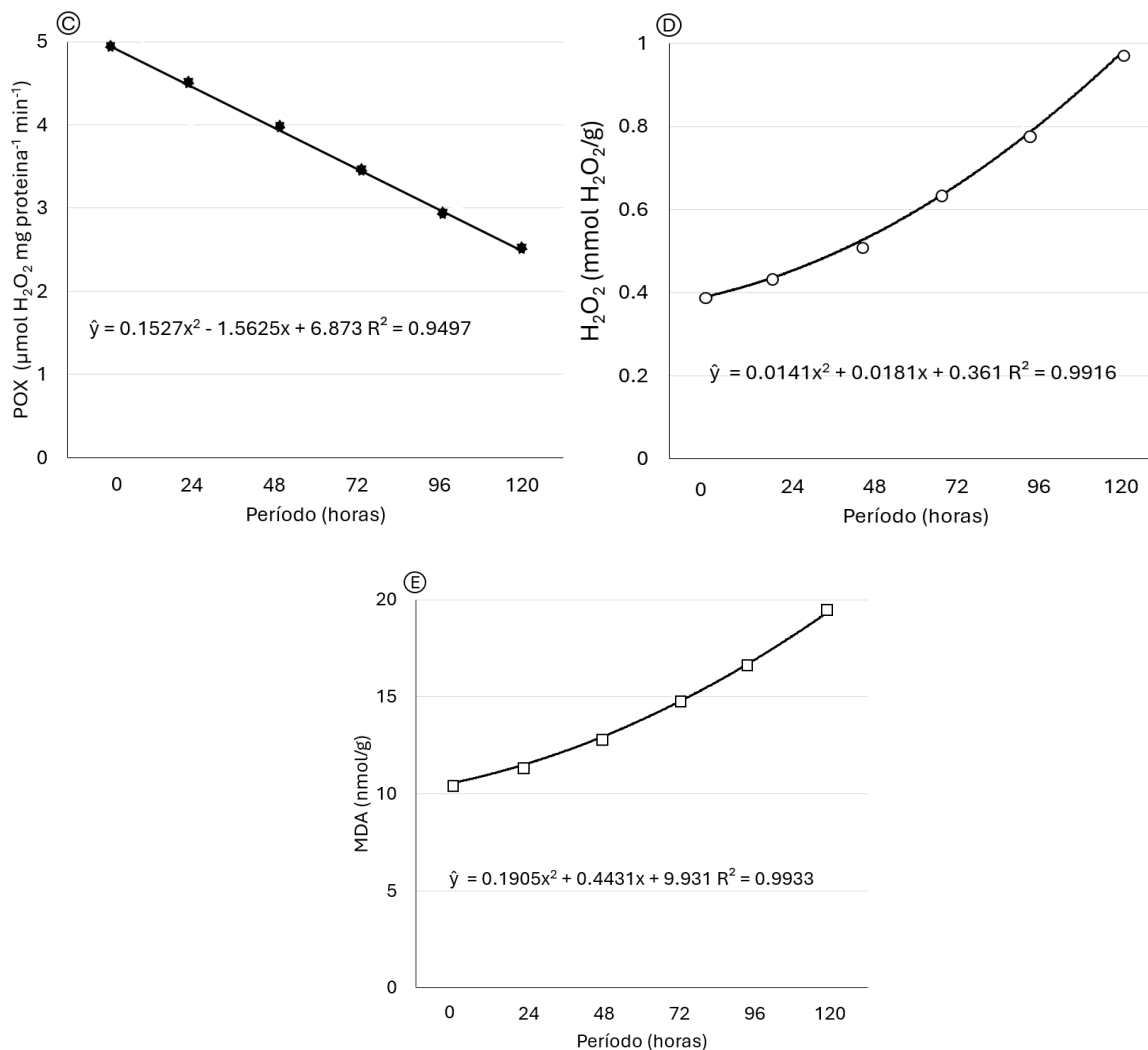


Figura 2. Atividade das enzimas antioxidantes CAT (A), SOD (B), POX (C) e ac\u00famulo de H₂O₂ (D) e MDA (E) em sementes de algod\u00e3o ap\u00f3s per\u00edodos de envelhecimento artificial.

Houve aumento no cont\u00e9udo de per\u00f3xido de hidrog\u00eanio (H₂O₂) e de malonalde\u00eddo (MDA) com o aumento no per\u00edodo de envelhecimento artificial (Figura 2D e 2E). Paralelamente, houve redu\u00e7\u00e3o no potencial fisiol\u00f3gico das sementes conforme indicam os resultados de germina\u00e7\u00e3o (Figura 1A) e dos testes de vigor empregados (Figuras 1B e 1C). Na Figura 3, observa-se que os coeficientes de correla\u00e7\u00e3o entre essas vari\u00e1veis bioqu\u00edmicas e as vari\u00e1veis relacionadas \u00e0 qualidade fisiol\u00f3gica das sementes foram negativos e superiores a 0,90.

\u00c9 importante destacar que o H₂O₂ \u00e9 uma das principais EROs formadas durante o metabolismo celular (Bailey, 2004). Em n\u00edveis moderados, o H₂O₂ atua como uma mol\u00e9cula sinalizadora regulando processos fisiol\u00f3gicos importantes como a germina\u00e7\u00e3o, quebra da dorm\u00eancia e a ativa\u00e7\u00e3o de mecanismos de defesa antioxidante (Bailey, 2004). No entanto,

durante o processo de deterioração, ocorre um acúmulo excessivo de H_2O_2 devido ao desequilíbrio entre sua produção e a capacidade de remoção pelas enzimas antioxidantes, como CAT, SOD e POX (Bailey, 2004; Zhou et al., 2019). Esse acúmulo provoca estresse oxidativo, resultando em danos que comprometem a integridade das membranas e reduz germinação e o vigor das sementes (Kibinza et al., 2011; Wang et al., 2012). Assim, o aumento dos níveis de H_2O_2 , aliado à redução das enzimas do sistema antioxidante (Figura 2A, 2B e 2C) observadas no nosso estudo evidenciam este desequilíbrio e nos auxiliam a compreender os resultados observados na germinação e no vigor por meio do envelhecimento artificial (Figura 1A, 1B e 1C) (Bhatia et al., 2010). Em sementes oleaginosas como a soja, Pinheiro et al. (2023) constataram que a deterioração por umidade na fase de pré-colheita acarretou o acúmulo de peróxido de hidrogênio nas sementes de diferentes cultivares. Em sementes de arroz submetidas ao envelhecimento acelerado, houve acúmulo significativo de H_2O_2 , intensificando o estresse oxidativo decorrente da redução da eficiência do sistema antioxidante. O excesso de espécies reativas de oxigênio promove peroxidação lipídica, comprometendo a integridade das membranas celulares, além de causar danos a proteínas, enzimas e ao DNA. Esses efeitos resultam em desorganização metabólica, redução da produção de energia e prejuízo ao desenvolvimento do embrião, culminando com redução da germinação (Zheng et al., 2024).

Já o MDA é um dos principais produtos da peroxidação lipídica, sendo utilizado como indicador bioquímico de dano oxidativo e de deterioração em sementes (Goel e Sheoran, 2003). Durante o envelhecimento natural ou artificial das sementes de algodão, há uma redução da atividade das enzimas antioxidantes, conforme observado nas Figuras 2A, 2B e 2C, o que resulta em acúmulo de EROs (Dhaliwal et al., 2024). Esse desequilíbrio oxidativo leva à oxidação dos lipídios de membrana, aumentando a permeabilidade celular e promovendo vazamento de solutos, como açúcares e eletrólitos que são sintomas típicos da perda de integridade de membranas (Goel e Sheoran, 2003b). Nessa etapa, o MDA se acumula como produto da degradação dos fosfolipídios, refletindo diretamente o grau de deterioração fisiológica (Sai et al., 2025). Em sementes de algodão, esse aumento de MDA tem sido fortemente associado à redução do vigor e da germinação, também observado em nossos estudos (Figura 1A, 1B e 1C). Estudos recentes mostram que, à medida que o tempo de envelhecimento artificial avança, os teores de MDA aumentam de forma proporcional à queda do potencial fisiológico, indicando que o composto é um marcador confiável de estresse oxidativo e perda de viabilidade (Sai et al., 2025; Pinheiro et al., 2023). Além disso, o acúmulo de MDA nas sementes pode afetar negativamente a integridade do embrião e o metabolismo

inicial da germinação, comprometendo no tamanho e na formação de plântulas normais (Goel e Sheoran, 2003a). Pesquisas realizadas por Wang et al. (2023) e Zhang et al. (2022) reforçam que o MDA pode ser utilizado em conjunto com indicadores como H₂O₂ e a atividade de enzimas antioxidante para caracterizar o grau de deterioração oxidativa em sementes de algodão submetidas ao armazenamento prolongado ou a condições de estresse térmico e hídrico. Assim, a quantificação do MDA está relacionada à qualidade fisiológica das sementes.

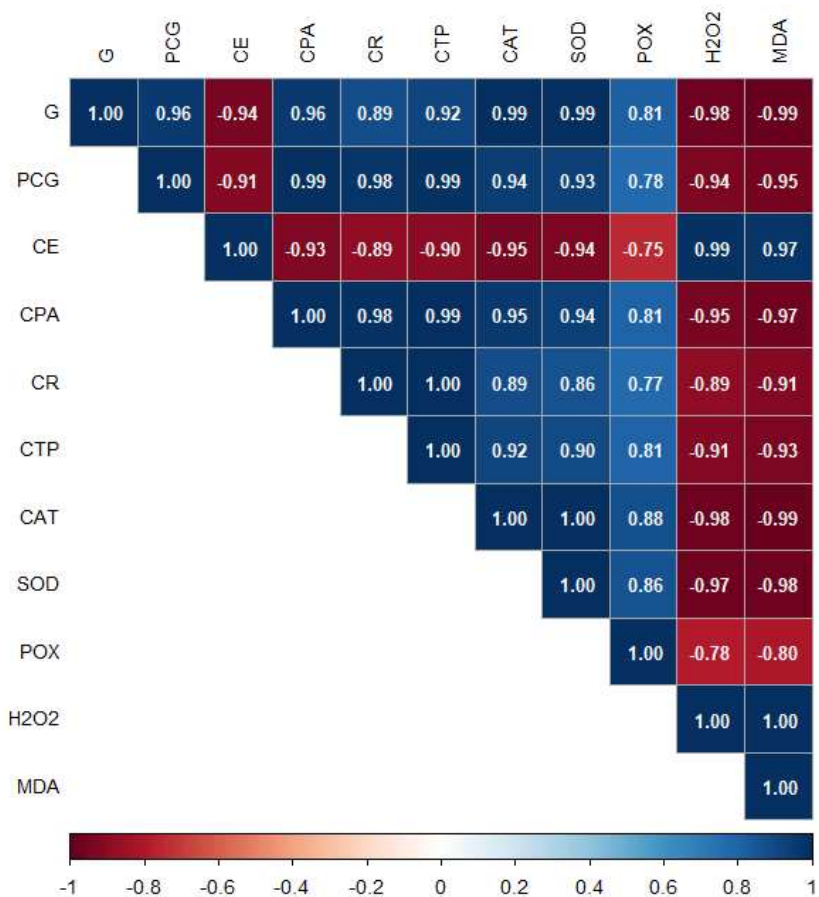


Figura 3. Análise de correlação de Pearson (r) das variáveis fisiológicas: germinação (G), primeira contagem de germinação (PCG), condutividade elétrica (CE), comprimento da parte aérea (CPA), comprimento de radícula (CR) e comprimento total de plântulas (CTP), e das análises bioquímicas catalase (CAT), superóxido dismutase (SOD) e peroxidases (POX), peróxido de hidrogênio (H₂O₂) e malonaldeído (MDA).

Espectroscopia no infravermelho próximo e análise de imagens multiespectrais

Diferentes métodos e algoritmos foram testados para o desenvolvimento do modelo de classificação das sementes de algodão submetidas ao envelhecimento artificial, tanto para a técnica FT-NIR como para a análise de imagens multiespectrais (Tabela 1). O método de PLS-

DA apresentou maior assertividade e concordância ao utilizar a segunda derivada Savitzky-Golay para a técnica FT-NIR, com valores de acurácia e kappa de 100%, diferentemente do que foi observado para a técnica de análise de imagens multiespectrais, em que todos os pré-processamentos apresentaram valores de acurácia inferior a 80% e kappa inferior a 65% (Tabela 1). O algoritmo Savitzky-Golay é uma técnica de suavização e cálculo de derivadas que é extremamente útil no processamento de dados espectrais. Ele não só ajuda a suavizar o sinal, removendo o ruído de alta frequência, mas também permite o cálculo das derivadas de primeira e segunda ordem, essa última remove a linha de base e a tendência linear, úteis para detectar mudanças rápidas e curvaturas nos dados (Savitzky e Golay, 1964). No nosso estudo, a segunda derivada mostrou-se maior assertividade (Tabela 1, Figura 4C). Soares et al. (2016) estudaram a classificação de variedades de sementes de algodão usando imagens hiperespectrais no infravermelho próximo, e concluíram que o modelo PLS-DA com a segunda derivada Savitzky-Golay permitiu taxas de classificação de 98,0%, 89,7% e 91,7% para a técnica NIR convencional e VIS-NIR convencional. O pré-processamento da segunda derivada Savitzky-Golay por meio da técnica de PLS-DA também foi eficiente para classificação de sementes de soja (Silva et al., 2024), lentilha (Limão et al., 2025), braquiária (Souza et al., 2023) e milho (Andrade et al., 2020; Andriazzi et al., 2023).

O método de Random Forest (Tabela 1) apresentou maiores taxas de acerto e concordância para a técnica de análise de imagens multiespectrais já com os dados originais, com alto valor de acurácia de 90% e kappa de 84%. Ao aplicar os métodos de pré-processamento SNV e MSV, houve aumento nos valores de acurácia (92%) e kappa (88%), o que não foi observado para a técnica do FT-NIR, uma vez que pelos dados originais verificou-se baixo valor de acurácia (50%) e kappa (26%); para todos os pré-processamentos utilizados foram obtidos valores de acurácia inferiores a 87% e kappa inferior a 80% (Tabela 1). O Random Forest é um método de aprendizado de máquina que se baseia na combinação de diferentes árvores de decisão para realizar classificação, gerando classificadores de forma aleatória e realiza combinação dos classificadores até produzir o resultado final (Santana et al., 2020).

Tabela 1. Valores de acurácia e coeficiente kappa dos diferentes métodos de pré-processamento obtidos pelos modelos de classificação PLS-DA, Random Forest e Neural Network para os os níveis de potencial fisiológico de sementes de algodão usando dados espectrais de FT-NIR e análise de imagens multiespectrais.

| Métodos | Pré-processamento | FT-NIR (n=120) | | Multiespectral (n=60) | |
|-------------------|----------------------|-------------------|-------|--------------------------|-------|
| | | Acurácia | Kappa | Acurácia | Kappa |
| Validação | | | | | |
| PLS-DA | Dados brutos | 0,77 | 0,66 | 0,75 | 0,62 |
| | SNV | 0,78 | 0,66 | 0,69 | 0,54 |
| | MSC | 0,78 | 0,66 | 0,63 | 0,44 |
| | 1st Derivative of SG | 0,98 | 0,97 | - | - |
| | 2nd Derivative of SG | 1,00 | 1,00 | - | - |
| Random Forest | Raw | 0,50 | 0,26 | 0,90 | 0,84 |
| | SNV | 0,72 | 0,58 | 0,92 | 0,88 |
| | MSC | 0,73 | 0,59 | 0,92 | 0,88 |
| | 1st Derivative of SG | 0,85 | 0,79 | - | - |
| | 2nd Derivative of SG | 0,86 | 0,79 | - | - |
| Neural Network | Raw | 0,76 | 0,65 | 0,87 | 0,76 |
| | SNV | 0,81 | 0,72 | 0,89 | 0,79 |
| | MSC | 0,82 | 0,72 | 0,90 | 0,81 |
| | 1st Derivative of SG | 0,94 | 0,83 | - | - |
| | 2nd Derivative of SG | 0,96 | 0,95 | - | - |

n = número de espectros usados para testar o modelo de classificação. SNV = Standard Normal Variate; MSC = Multiplicative Scatter Correction; SG = Savitzky-Golay.

Quando aplicado com as técnicas de pré-processamentos eficientes, o Random Forest permite maior assertividade dos resultados, conforme observado na tabela 1 para os dados multiespectrais, onde os pré-processamentos SNV e MSC apresentaram altos valores de acurácia e kappa. Os métodos de correção da dispersão como o SNV e MSC tem o objetivo de reduzir a variabilidade física entre as amostras devido a dispersão (Delwiche, 1998). O MSC, por exemplo é muito poderosa para corrigir a dispersão multiplicativa dos dados espectrais; já o SNV é eficiente para normalizar os dados, visando remover variações de amplitudes ou efeitos de dispersão (Delwiche, 1998). O Random Forest têm sido cada vez mais aplicado em modelagem e predição da qualidade de sementes conforme observado em estudos para soja (Medeiros et al., 2020a; Barboza et al., 2021) milho (Andriuzzi et al., 2023), trigo (Singh et al., 2020), e braquiária (Medeiros et al., 2020b).

Já o modelo Redes Neurais apresentou um bom desempenho tanto para a técnica do FT-NIR como para a de análise de imagens multiespectrais (Tabela 1). Para o FT-NIR, ao utilizar o pré-processamento pela segunda derivada de Savitzky-Golay o valor de acurácia foi de 96% e kappa de 95%, mostrando alta assertividade do modelo. (Tabela 1). Já para a técnica de análise de imagens multiespectrais, obteve-se acurácia de 90% e kappa de 81%, quando se utilizou pré-processamento pelo MSC, que mostrando-se também eficiente para a classificação do modelo, mas com desempenho inferior ao observado pelo método de Random Forest (Tabela 1). O modelo baseado em Redes Neurais também se mostrou eficiente na classificação de danos de sementes de soja, chegando em valores de acurácia de 100% ao classificar sementes sadias e daquelas danificadas (Wang, Ram e Dowell, 2002), como também na classificação de diferentes espécies de trigo (Nabeel e Quteishat, 2015) e milho (Andriazzi et al., 2023).

Os espectros originais (Figura 4A) e a média dos espectros (Figura 4B) nos mostram que houve diferença na reflectância das sementes com diferentes níveis de potencial fisiológico. Conforme citado anteriormente, o pré-processamento utilizando a segunda derivada Savitzky-Golay pelo método de PLS-DA foi o mais eficiente para o desenvolvimento do modelo para a técnica do FT-NIR e essa assertividade pode ser comprovada por meio da matriz de confusão ilustrada na figura 4C, onde é possível destacar assertividade de 100% para as três classes de potencial fisiológico (alta, média e baixa).

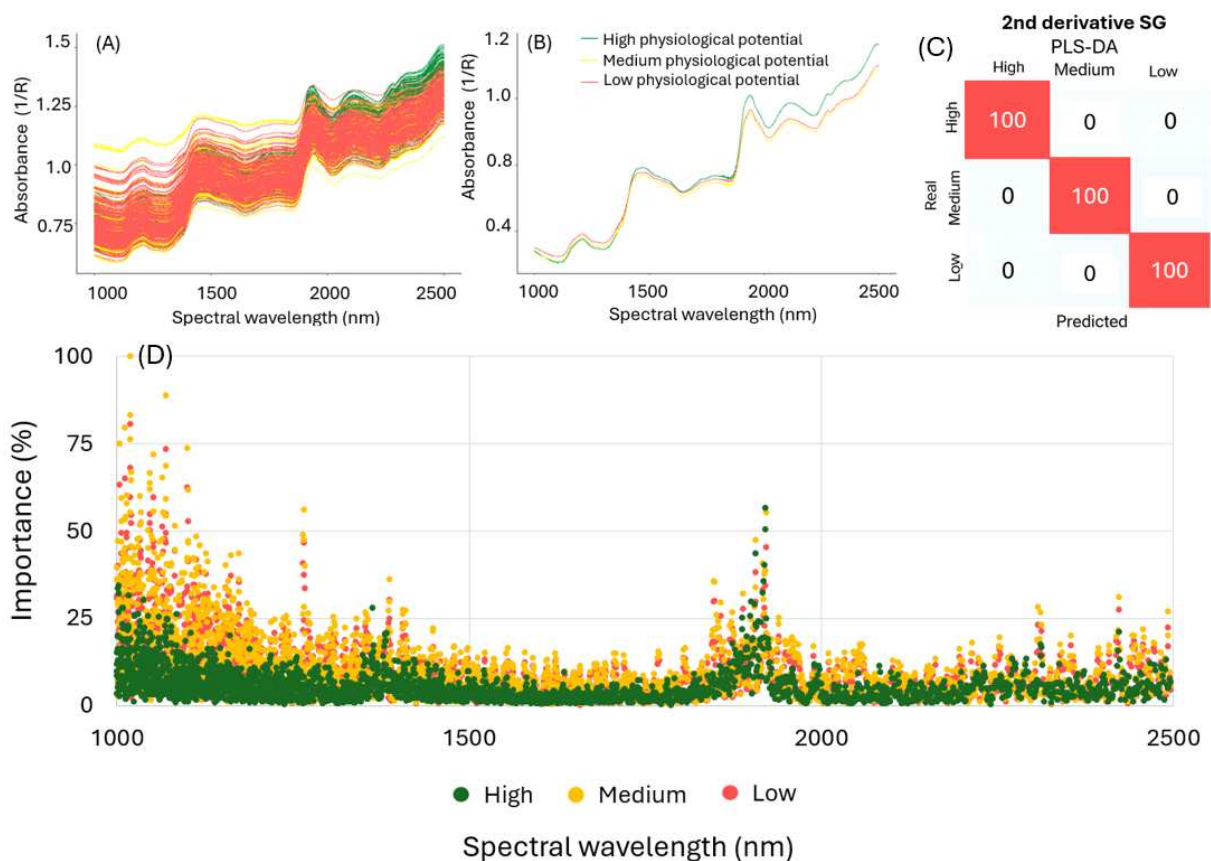


Figura 4. Espectros originais (A) e média dos espectros (B); matriz de confusão dos dados de validação utilizando o modelo PLS-DA com o pré-processamento da 2ª derivada Savitzky-Golay (C); importância das variáveis dos comprimentos de onda que mais contribuíram para o desenvolvimento do modelo PLS-DA (D) de acordo com os níveis de potencial fisiológico de sementes de algodão pela técnica FT-NIR.

Com relação aos comprimentos de onda que mais contribuíram para o desenvolvimento do modelo PLS-DA para a técnica do FT-NIR (Figura 4D), destaca-se a faixa de 1000-1100 nm, 1300 nm e 1900 nm. A faixa de 1.000-1.100 nm corresponde predominante aos grupos das hidroxilas (O-H) associados ao conteúdo de carboidratos (Workman et al., 2007). A faixa de 1.300 nm compreende os grupos funcionais carboxílicos e hidroxílicos relacionados a ácidos graxos mono e poli-insaturados e saturados (C-H) (Ozaki et al., 2006), e também a proteínas (O-H) (Ambrose et al., 2016; Xu et al., 2020). Já o comprimento de 1.900 nm apresenta relação com o conteúdo de proteínas, carboidratos e água (Kusumaningrum et al. 2018; Carreiro Soarres et al., 2016). Dessa forma, as variações observadas nos espectros, segundo as classes de potencial fisiológico a partir dos períodos de envelhecimento artificial, indicam alterações na composição química das sementes, especialmente em lipídios, carboidratos e proteínas.

Esses compostos representam as principais reservas das sementes de algodão, que são classificadas como oleaginosas, tornando-as mais suscetíveis à deterioração ao longo das diferentes fases do processo produtivo (Mata et al., 2022). Assim, a intensidade das mudanças bioquímicas causadas pela deterioração reflete diretamente no potencial fisiológico das sementes (Marcos-Filho, 2015).

Em síntese, nosso estudo comprovou a eficiência da espectroscopia no infravermelho próximo para a separação dos diferentes níveis de potencial fisiológico de sementes de algodão, principalmente quando utilizado com junto com algoritmos de aprendizado de máquina (Tabela 1, Figura 4A, 4B, 4C e 4D), que possibilitam realizar a modelagem correta e a relação dos dados espectrais com a composição química das sementes (Medeiros et al., 2020a; Burns e Ciurczak, 2007). Essa técnica vem sendo amplamente estudada em diversos segmentos alimentícios, industriais, farmacêuticos e na agricultura (Quentin et al., 2016; Samamad et al., 2018; Han et al., 2021; Liu et al., 2021). Em sementes, os primeiros estudos utilizando essa técnica foram realizados para a determinação do teor de água de sementes de soja e milho por Hart, Norris e Golumbic, (1962). Os recentes avanços utilizando essa técnica confirmam que é possível aplicá-la na separação de genótipos de soja (Soares et al., 2024), milho (Schutz et al., 2022), braquiária (Souza et al., 2023), grão-de-bico (Ribeiro et al., 2021). Na separação de sementes envelhecidas artificialmente de melancia (Yasmin et al., 2019) e lentilha (Limão et al., 2025). O NIR também é promissor para avaliar a composição química de proteínas, carboidratos e óleos de sementes de algodão (Zhuang et al., 2023), ervilha (Hacisalihoglu et al., 2020), soja (Ferreira et al., 2014), trigo, aveia e cevada (Baianu e Guo, 2011).

Para a técnica de análise de imagens multiespectrais, a figura 5A mostra a escala de cinza e a média da escala de cinza (Figura 5B) conforme as classes de potencial fisiológico das sementes de algodão. A matriz de confusão para o método de Random Forest por meio do pré-processamento MSC é apresentada na figura 5C, nela destaca-se os acertos e a distribuição dos erros conforme as classes de potencial fisiológico, em que foi possível observar assertividade de 100% para a classe de alto, 90% para médio e 88% para a classe de baixo potencial fisiológico, conseqüentemente, com pouca distribuição de erros das classes.

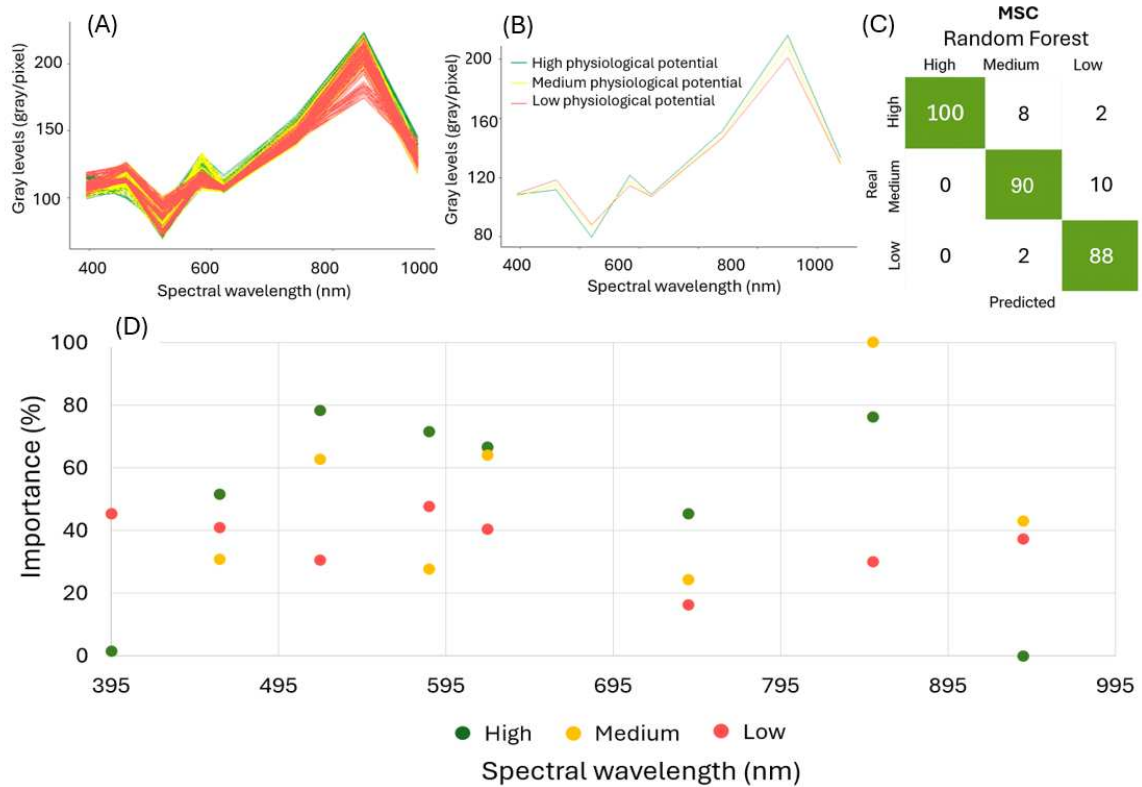


Figura 5. Escala de cinza com os dados originais (A) e média da escala de cinza (B); matriz de confusão dos dados de validação utilizando o modelo Random Forest com o pré-processamento MSC (C); importância das variáveis dos comprimentos de onda que mais contribuíram para o desenvolvimento do modelo Random Forest (D) de acordo com os níveis de potencial fisiológico de sementes de algodão pela técnica de análise de imagens multiespectrais.

Dentre os comprimentos de onda que mais contribuíram para o desenvolvimento do modelo por meio da técnica de análise de imagens multiespectrais, destaca-se 480 nm, 520 nm, 580 nm, 610 nm, 740 nm e 850 nm (Figura 5D). Os comprimentos de onda de 480 nm e 520 nm (região azul) correspondem a regiões sensíveis a pigmentos naturais no tegumento da semente, como fenóis oxidados, flavonoides e clorofilas residuais, os quais são importantes para diferenciar a qualidade fisiológica, pois alterações químicas durante a deterioração podem afetar a absorção/reflexão nessa faixa de comprimento de onda (Liu et al., 2019). A faixa de comprimento de onda de 610 nm (região vermelha) está relacionado à absorção de clorofila a e b (Liu et al., 2019), 740 nm (região próxima ao vermelho) corresponde à transição entre forte absorção de vermelho e níveis médios de cinza mais altos. Nessa faixa, a reflectância aumenta à medida que a semente perde clorofila; assim, sementes com maior potencial fisiológico, integridade e viabilidade apresentam valores de reflectância mais altos nessa região (Gitelson et al. 2003). Vale ressaltar também que essa faixa é sensível ao teor de água da semente

(Gitelson et al. 2003). Já o comprimento de onda na faixa de 840-850 nm é relacionada ao teor de água e conteúdo de lipídeos (Gitelson et al. 2003). A integração de imagens multiespectrais com técnicas quimiométricas possibilita a obtenção de informações detalhadas sobre múltiplos aspectos das sementes, como suas características morfológicas, químicas e fisiológicas (Petronilio et al., 2025). Esse tipo de abordagem tem ganhado destaque entre pesquisadores do mundo todo, por permitir análises rápidas, precisas e não invasivas, representando uma alternativa eficiente ou complementar aos métodos convencionais de avaliação da qualidade de sementes.

No nosso estudo, a análise de imagens multiespectrais, utilizando um protótipo, combinada com algoritmos de aprendizado de máquina também se mostrou eficiente e promissora na separação dos diferentes níveis de potencial fisiológico de sementes de algodão (Tabela 1, Figura 5A, 5B, 5C e 5D). Essa técnica vêm sendo objeto de estudo no desenvolvimento de programas de controle de qualidade de sementes (Petronilio et al., 2025). A eficiência dessa técnica já foi comprovada em sementes de soja envelhecidas artificialmente para obtenção de sementes com diferentes níveis de qualidade fisiológica, sendo usada para verificar as alterações ópticas dos tecidos das sementes e classificar sementes de alto e baixo vigor (Barboza da Silva et al., 2021), para a identificação de sementes de soja produzidas sob condições favoráveis e desfavoráveis (Petronilio et al., 2025), para classificação de sementes e melancia (Liu et al., 2019) e trigo (Vresak et al., 2016) quanto ao potencial fisiológico, como também na detecção de danos em sementes de milho (Sendin, Manley e Williams, 2018).

Nos últimos anos, a aplicação combinada de diferentes técnicas para classificar sementes tem se intensificado no campo da análise e tecnologia de sementes. Nosso estudo de deterioração de sementes de algodão, por meio da técnica de envelhecimento artificial, permitiu observar alterações bioquímicas e fisiológicas e realizar a classificação com base no potencial fisiológico por meio de técnicas recentes de espectroscopia NIR, análise de imagens multiespectrais e algoritmos de aprendizado de máquina. Os resultados obtidos neste estudo ajudam a elucidar a importância da aplicação de tecnologias emergentes rápidas e não destrutivas relacionadas à composição química das sementes, destacando-se como um avanço tecnológico para as empresas produtoras de sementes em programas de gestão da qualidade.

CONCLUSÕES

A germinação e o vigor das sementes de algodão bem como a atividade das enzimas do sistema antioxidantes SOD, CAT e POX decrescem com o aumento da deterioração ocasionada pelo envelhecimento artificial, ocorrendo aumento no acúmulo de H₂O₂ e MDA.

A espectroscopia no infravermelho próximo (FT-NIR) é eficiente para classificar sementes de algodão de acordo com seu nível de deterioração, com destaque para o algoritmo PLS-DA e o pré-processamento com a 2ª derivada Savitzky-Golay, alcançando acurácia de 100%.

O modelo baseado em Random Forest, por meio do pré-processamento MSC, mostrou-se eficiente para classificar sementes de algodão quanto ao nível de deterioração pela técnica de análise de imagens multiespectrais, alcançando acurácia de até 92%.

CONCLUSÕES GERAIS

A técnica de Raios X apresenta relação com o potencial fisiológico das sementes de algodão, além de fornecer informações sobre suas estruturas internas.

Aos três dias de germinação, o comprimento da parte aérea, da raiz primária, o comprimento total de plântula e os índices de uniformidade, crescimento e vigor, determinados pelo Groundeye, são eficientes na classificação de lotes de algodão quanto ao potencial fisiológico.

O aprendizado de máquina supervisionado por meio do Ilastik possibilitou a classificação dos lotes quanto ao potencial fisiológico, obtendo resultados comparáveis aos testes convencionais.

A espectroscopia no infravermelho próximo (FT-NIR) é eficiente para a classificação de sementes de algodão de acordo com seu potencial fisiológico. Todos os modelos testados mostraram-se promissores e precisos, com destaque para o PLS-DA, especialmente quando utilizado o pré-processamento com a 2ª derivada de Savitzky-Golay, alcançando acurácia de até 98%. O modelo Random Forest, utilizando os dados brutos, foi o mais eficiente na classificação do potencial fisiológico de sementes de algodão por meio de imagens multiespectrais, atingindo acurácia de até 81%.

Esses métodos são promissores para a classificação rápida, não destrutiva e precisa da qualidade de sementes de algodão em programas de controle de qualidade de sementes.

A germinação e o vigor das sementes de algodão bem como a atividade das enzimas do sistema antioxidantes SOD, CAT e POX decrescem com o aumento da deterioração ocasionada pelo envelhecimento artificial, ocorrendo aumento no acúmulo de H₂O₂ e MDA .

A espectroscopia no infravermelho próximo (FT-NIR) é eficiente para classificar sementes de algodão de acordo com seu nível de deterioração, com destaque para o algoritmo PLS-DA e o pré-processamento com a 2ª derivada Savitzky-Golay, alcançando acurácia de 100%.

O modelo baseado em Random Forest, por meio do pré-processamento MSC, mostrou-se eficiente para classificar sementes de algodão quanto ao nível de deterioração pela técnica de análise de imagens multiespectrais, alcançando acurácia de até 92%.

REFERÊNCIAS

ABRAPA. Evolução da balança comercial anual de algodão do Brasil e UFs. 2025b. Disponível em: <https://app.powerbi.com/view?r=eyJrIjoiMzlmMzNiMjltMjg4Mi00NWYyLWJhYzQtMmVlMmI4OGQ5YjU0IiwidCI6IjRhMDk1OGIzLTg4MWQtNDBmYS05NTU1LTlwODQ1MzdhYWZkMyJ9>. Acesso em: 5 nov. 2025.

AMBROSE, A.; KANDPAL, L. M.; KIM, M. S.; LEE, W. H.; CHO, B. K. High speed measurement of corn seed viability using hyperspectral imaging. *Infrared Physics e Technology*, v. 75, p. 173-179, 2016. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S1350449515300074>. Acesso em: 28 out. 2025.

ANDERSON, M. D.; PRASAD, T. K.; STEWART, C. R. Changes in isozyme profiles of catalase, peroxidase, and glutathione reductase during acclimation to chilling in mesocotylus of maize seedlings. *Plant Physiology*, v. 109, p. 1247-1257, 1995.

ANDRADE, G. C.; COELHO, C. M. M.; UARROTA, V. G. Modelling the vigour of maize seeds submitted to artificial accelerated ageing based on ATR-FTIR data and chemometric tools (PCA, HCA and PLS-DA). *Heliyon*, v. 6, n. 2, p. e03477, 2020. DOI: 10.1016/j.heliyon.2020.e03477.

ANDRIAZZI, C. V. G.; ROCHA, D. K.; CUSTÓDIO, C. C. Determination of the physiological quality of corn seeds by infrared equipment. *Journal of Seed Science*, 45, 2023. <https://doi.org/10.1590/2317-1545v45265346>

ARAÚJO, J. O.; DIAS, D. C. F. S.; NASCIMENTO, W. M.; MARTINS, A. O.; LIMÃO, M. A. R. Accelerated aging test and antioxidant enzyme activity to assess chickpea seed vigor. *Journal of Seed Science*, v. 43, n. 1, e202143038, 2021. DOI: 10.1590/2317-1545v43253934.

BAIANU, I.; GUO, J. NIR Calibrations for Soybean Seeds and Soy Food Composition Analysis: Total Carbohydrates, Oil, Proteins and Water Contents. *Nature Precedings*, 16 nov. 2011.

BAILEY, C. Active oxygen species and antioxidants in seed biology. *Seed Science Research*, v. 14, n. 2, p. 93-107, 2004.

BAILEY, C. ROS in seed germination. *Advances in Botanical Research*, v. 105, p. 177-204, 2023.

BARBOZA DA SILVA, C.; OLIVEIRA, N. M.; DE CARVALHO, M. E. A.; DE MEDEIROS, A. D.; DE LIMA NOGUEIRA, M.; DOS REIS, A. R. Autofluorescence-spectral imaging as an innovative method for rapid, non-destructive and reliable assessing of soybean seed quality. *Scientific Reports*, v. 11, p. 17834, 2021. Disponível em: <https://doi.org/10.1038/s41598-021-97223-5>.

BEAUCHAMP, C.; FRIDOVICH, I. Superoxide dismutase: improved assays and an assay applicable to acrylamide gels. *Analytical Biochemistry*, v. 44, n. 1, p. 276-287, 1971.

BHATIA, V. S. et al. Field deterioration of soybean seed: role of oxidative stresses and antioxidant defense mechanism. *Journal of Plant Biology*, v. 32, n. 2, p. 179-190, 2010.

BORBA, I. C. G.; BANDEIRA, J. M.; MARINI, P.; MARTINS, A. B. N.; MORAES, D. M. Metabolismo antioxidativo para separação de lotes de sementes de diferentes graus de homogeneidade. *Revista Brasileira de Biociências*, v. 12, n. 1, p. 20-26, 2014.

BRADFORD, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, v.72, p.248-254, 1976.

BRASIL. Ministério da Agricultura e Pecuária. Secretaria de Defesa Agropecuária. *Regras para Análise de Sementes – RAS*. Brasília: MAPA/SDA, 2025. https://wikisda.agricultura.gov.br/pt-br/Laborat%C3%B3rios/Metodologia/Sementes/RAS_2025/cap_15_tol_fiscalizacao

BROCHADO, M. G. S.; DE NORONHA, B. G.; DA COSTA LIMA, A.; GUEDES, A. G.; DA SILVA, R. C.; DOS SANTOS DIAS, D. C. F. et al. What is the most effective analytical method

for quantification and identification of microplastics in contaminated soils? *Environmental Geochemistry and Health*, v. 46, n. 7, p. 260, 2024.

BURNS, D. A.; CIURCZAK, E. W. *Handbook of Near-Infrared Analysis* (D. A. Burns e E. W. Ciurczak, 3 eds.). Boca Raton: CRC Press, p. 834, 2007.

CAKMAK, I.; HORST, W. J. Effect of aluminium on lipid peroxidation, superoxide dismutase, catalase, and peroxidase activities in root tips of soybean (*Glycine max*). *Physiologia Plantarum*, v. 83, n. 3, p. 463-468, 1991.

CAMPOS, Leonardo Vieira; RODRIGUES, Arthur Almeida; SALES, Juliana de Fátima; RODRIGUES, Douglas Almeida; VASCONCELOS FILHO, Sebastião Carvalho; RODRIGUES, Cássia Lino; VIEIRA, Dheyne Alves; CASTRO, Stella Tomaz de; NETO, Aurélio Rubio. Radiographic imaging as a quality index proxy for *Brachiaria brizantha* seeds. *Plants*, Basel, v. 11, n. 8, p. 1014, 2022. DOI: <https://doi.org/10.3390/plants11081014>.

CARREIRO SOARES, S. F.; MEDEIROS, E. P.; PASQUINI, C.; DE LELIS MORELLO, C.; HARROP GALVÃO, R. K.; UGULINO ARAÚJO, M. C. Classification of individual cotton seeds with respect to variety using near-infrared hyperspectral imaging. *Analytical Methods*, v. 8, n. 48, p. 8498-8505, 2016.

CEMBROWSKA-LECH, D.; RYBAK, K. Nanoprimering of barley seeds—a shotgun approach to improve germination under salt stress conditions by regulating of reactive oxygen species. *Plants*, v. 12, n. 2, 405, 2023.

CHANCE, B.; MAEHLY, A. C. Assay of catalases and peroxidases. *Methods in Enzymology*, v. 2, p. 764-775, 1955.

DELOUCHE, J. C.; BASKIN, C. C. Accelerated aging techniques for predicting the relative storability of seed lots. *Seed Science and Technology*, v. 1, n. 2, p. 427-452, 1973.

DELWICHE, S. R. Protein content of single kernels of wheat by near-infrared reflectance spectroscopy. *Journal of Cereal Science*, v. 27, n. 3, p. 241-254, 1998.

DHALIWAL, L. K.; SHIM, J.; ANGELES-SHIM, R. Fatty acid unsaturation improves germination of upland cotton (*Gossypium hirsutum*) under cold stress. *Frontiers in Plant Science*, v. 15, p. 1286908, 2024. DOI: [10.3389/fpls.2024.1286908](https://doi.org/10.3389/fpls.2024.1286908).

FERREIRA, D. S. et al. Comparison and application of near-infrared (NIR) and mid-infrared (MIR) spectroscopy for determination of quality parameters in soybean samples. *Food Control*, v. 35, n. 1, p. 227–232, 2014. DOI: 10.1016/j.foodcont.2013.07.010

FRANÇA-SILVA, F.; GOMES JUNIOR, F. G.; REGO, C. H. Q.; MARASSI, A. G.; TANNÚS, A. Advances in imaging technologies for soybean seed analysis. *Journal of Seed Science*, v. 45, e202345022, 2023. DOI: 10.1590/2317-1545v45274098.

GALLETTI, P. A.; CARVALHO, M. E. A.; HIRAI, W. Y.; BRANCAGLIONI, V. A.; ARTHUR, V.; SILVA, C. B. DA. Integrating optical imaging tools for rapid and non invasive characterization of seed quality: Tomato (*Solanum lycopersicum* L.) and Carrot (*Daucus carota* L.) as study cases. *Frontiers in Plant Science*, v. 11, Article 577851, 2020. <https://doi.org/10.3389/fpls.2020.577851>

GAY, C.; GEBICKI, J. M. A critical evaluation of the effect of sorbitol on the ferric–xylenol orange hydroperoxide assay. *Analytical Biochemistry*, v. 284, n. 2, p. 217-220, 2000.

GILL, S. S.; ANJUM, N. A.; GILL, R.; YADAV, S.; HASANUZZAMAN, M.; FUJITA, M.; MISHRA, P.; SABAT, S. C.; TUTEJA, N. Superoxide dismutase — mentor of abiotic stress tolerance in crop plants. *Environmental Science and Pollution Research*, v. 22, n. 14, p. 10375-10394, 2015.

GITELSON, A. A.; GRITZ, Y.; MERZLYAK, M. N. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, v. 160, n. 3, p. 271-282, 2003.

GOEL, A.; GOEL, A. K.; SHEORAN, I. S. Changes in oxidative stress enzymes during artificial ageing in cotton (*Gossypium hirsutum* L.) seeds. *Journal of Plant Physiology*, v. 160, p. 1093-1100, 2002. DOI: 10.1078/0176-1617-00881

GOEL, A.; SHEORAN, I. S. Lipid peroxidation and peroxide-scavenging enzymes in cotton seeds under natural ageing. *Biologia Plantarum*, v. 46, p. 429-434, 2003a. DOI: 10.1023/B:BIOP.0000024330.86124.45

GOEL, A.; SHEORAN, I. S. Lipid peroxidation and peroxide-scavenging enzymes in cotton seeds under natural ageing. *Biologia Plantarum*, v. 46, p. 429-434, 2003b. DOI: 10.1023/B:BIOP.0000024330.86124.45

HACISALIHOGU, G.; FREEMAN, J.; ARMSTRONG, P. R.; SEABOURN, B. W.; PORTER, L. D.; GUSTIN, J. F. Previsão de proteína, peso e óleo por espectroscopia de infravermelho próximo de semente única para seleção de qualidade de sementes e características de rendimento em ervilha (*Pisum sativum*). *J Sci Alimentos Agrícolas*, v. 100, p. 3488–3497, 2020. DOI: 10.1002/jsfa.10389

HAN, H.; WANG, Z.; ZHANG, Z.; LIU, X.; LI, H.; SONG, G.; WANG, C. Prediction of peanut kernel oil yield using near-infrared reflectance spectroscopy (In Chinese). *Shandong Agric. Ciência*, v. 53, 2021. DOI: 10.14083/j.issn.1001-4942.2021.11.020

HART, J. R.; NORRIS, K. H.; GOLUMBIC, C. Determination of the moisture content of seeds by near-infrared spectrophotometry of their methanol extracts. *Cereal Chemistry*, v. 39, p. 94-99, 1962. Disponível em: <https://www.cerealsgrains.org/publications/cc/backissues/1962/Documents/CC1962a11.html>

KIBINZA, S.; BAZIN, J.; BAILLY, C.; FARRANT, J. M.; CORBINEAU, F. Catalase is a key enzyme in seed recovery from ageing during priming. *Plant Science*, v. 181, p. 309-315, 2011.

KRZYZANOWSKI, F. C.; DIAS, D. C. F. S.; FRANÇA-NETO, J. F. B. Deterioração e vigor da semente. Circular Técnica 191. Londrina, PR, ISSN 2176-2864, 2022.

KRZYZANOWSKI, F. C.; FRANÇA-NETO, J. B.; GOMES-JUNIOR, F. G.; NAKAGAWA, J. Testes de vigor baseados em desempenho de plântulas. In: Krzyzanowski, F. C.; Vieira, R. D.; Marcos-Filho, J.; França Neto, J. B. (Eds.). *Vigor de sementes: conceitos e testes*. Londrina: ABRATES, 2020. p. 80-126.

KUHN, M. Building predictive models in R using the caret package. *Journal of Statistical Software*, v. 28, n. 5, 2008.

KUO, M. C.; KAO, C. H. Aluminum effects on lipid peroxidation and antioxidative enzyme activities in rice leaves. *Biologia Plantarum*, v. 46, n. ---

KUSUMANINGRUM, D.; LEE, H.; LOHUMI, S.; MO, C.; KIM, M. S.; CHO, B. Non-destructive technique for determining the viability of soybean (*Glycine max*) seeds using FT-NIR spectroscopy. *Journal of the Science of Food and Agriculture*, v. 98, n. 5, p. 1734-1742, 2018.

LIMÃO, M. A. R., PINHEIRO, D. T., MEDEIROS, A. D., RIBEIRO, J. P. O., DIAS, D. C. F. S., e SILVA, L. J.. Classification of artificially aged lentil seeds using Fourier transform near-

infrared spectroscopy. *Pesquisa Agropecuária Tropical*, 55, e82087, 2025b. <https://doi.org/10.1590/1983-40632025v5582087>

LIMÃO, M. A. R.; DIAS, D. C. F. S.; NASCIMENTO, W. M.; NORONHA, B. G.; SOARES, J. M.; SILVA, L. J. DA. Near-infrared spectroscopy and accelerated ageing in evaluating the vigour of lentil seeds. *Revista Ciência Agronômica*, v. 56, e202493547, 2025a. DOI: 10.5935/1806-6690.20250034

LIMÃO, M. A. R.; DIAS, D. C. F. S.; PERIS, G. C. O.; TRANCOSO, A. C. R.; MIRANDA, R. M.; NASCIMENTO, W. M. Adequação do teste de condutividade elétrica e atividade de enzimas antioxidantes para avaliar o vigor de sementes de lentilha (*Lens culinaris Medik.*). *Semina: Ciências Agrárias*, Londrina, v. 45, n. 3, p. 799-818, 2024. DOI: 10.5433/1679-0359.2024v45n3p799

LIU, C. ET AL. Nondestructive determination of transgenic *Bacillus thuringiensis* rice seeds (*Oryza sativa L.*) using multispectral imaging and chemometric methods. *Food Chemistry*, v. 153, p. 87–93, 2014.

LIU, W.; XU, X.; LIU, C.; ZHENG, L. Rapid discrimination of high-quality watermelon seeds by multispectral imaging combined with chemometric methods. *Journal of Applied Spectroscopy*, v. 85, n. 6, p. 1044-1049, 2019.

MARCOS-FILHO, J. *Fisiologia de sementes de plantas cultivadas*. 2. ed. Londrina: ABRATES, 2015. 659 p.

MARCOS-FILHO, J. Teste de envelhecimento acelerado. In: Krzyzanowski, F. C.; Vieira, R. D.; França-Neto, J. B.; Marcos-Filho, J. (Eds.). *Vigor de sementes: conceitos e testes*. Londrina: ABRATES, 2020. p. 182-244.

MATA, M. M.; ROCHA, P. D.; FARIAS, I. K. T.; SILVA, J. L. B.; MEDEIROS, E. P.; SILVA, C. S.; SIMÕES, S. S. Distinguishing cotton seed genotypes by means of vibrational spectroscopic methods (NIR and Raman) and chemometrics. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, v. 266, p. 120399, 2022. DOI: 10.1016/j.saa.2021.120399

MAYRINCK, L. G. et al. Use of near infrared spectroscopy in cotton seeds physiological quality evaluation. *Journal of Seed Science*, v. 42, p. 1-11, 2020. DOI: 10.1590/2317-1545v42227169.

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, n. 1, p. 11267, 2020a.

MEDEIROS, A. D.; SILVA, L. J.; RIBEIRO, J. P. O.; FERREIRA, K. C.; ROSAS, J. T. F.; SANTOS, A. A.; SILVA, C. B. Machine learning for seed quality classification: an advanced approach using merger data from FT-NIR spectroscopy and X-ray imaging. *Sensors*, v. 20, n. 15, p. 4319, 2020b. DOI: 10.3390/s20154319.

MITTLER, R. Oxidative stress, antioxidants and stress tolerance. *Trends in Plant Science*, v. 7, p. 405–410, 2002.

MITTLER, R. ROS are good. *Trends in Plant Science*, v. 22, n. 1, p. 11–19, jan. 2017. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S1360138516301121>.

NABEEL, A.; QUTEISHAT, A. Wheat seeds classification using multi-layer perceptron artificial neural network. *International Journal of Electronics Communication and Computer Engineering*, v. 6, p. 2278-4209, 2015.

OTSU, N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, v. 9, p. 62-66, 1979. DOI: 10.1109/TSMC.1979.4310076.

OZAKI, Y.; MCCLURE, F.; CHRISTY, A. Near-infrared spectroscopy in food science and technology. 1. ed. New Jersey: Wiley-Interscience, 2006. 406 p.

PARIDA, A. K.; DAS, A. B. Salt tolerance and salinity effects on plants: a review. *Ecotoxicology and Environmental Safety*, v. 60, n. 3, p. 324–349, 2005.

PEIXOTO, P. H. P.; CAMBRAIA, J.; SANT'ANNA, R.; MOSQUIM, P. R.; MOREIRA, M. A. Aluminum effects on lipid peroxidation and on the activities of enzymes of oxidative metabolism in sorghum. *Revista Brasileira de Fisiologia Vegetal*, v. 11, n. 3, p. 137-143, 1999.

PETRONILIO, A. C. P.; Mastrangelo, C. B.; Batista, T. B.; de Oliveira, G. R. F.; dos Santos, I. L.; da Silva, E. A. A. Smart and accurate: a new tool to identify stressed soybean seeds based on multispectral images and machine learning models. *Smart Agricultural Technology*, v. 12, p. 101042, 2025.

PINHEIRO, D. T.; DIAS, D. C. F. S.; SILVA, L. J. DA; MARTINS, M. S.; FINGER, F. L. Oxidative stress, protein metabolism, and physiological potential of soybean seeds under

weathering deterioration in the pre-harvest phase. *Acta Scientiarum. Agronomy*, v. 45, e56910, 2023. DOI: 10.4025/actasciagron.v45i1.56910.

QUENTIN, A. G. et al. Application of near-infrared spectroscopy for estimation of non-structural carbohydrates in foliar samples of *Eucalyptus globulus* Labillardière. *Tree Physiology*, v. 37, p. 131-141, 2016. DOI: 10.1093/treephys/tpw083.

R CORE TEAM. R: a language and environment for statistical computing. 2024. Disponível em: <https://www.R-project.org/>. Acesso em: 28 out. 2025.

RAY, P. D.; HUANG, B. W.; TSUJI, Y. Reactive oxygen species (ROS) homeostasis and redox regulation in cellular signaling. *Cellular Signaling*, v. 24, n. 2012, p. 981–990, 2012.

RIBEIRO, J. P. O.; MEDEIROS, A. D.; CALIARI, I. P.; TRANCOSO, A. C. R.; MIRANDA, R. M.; FREITAS, F. C. L.; SILVA, L. J.; DIAS, D. C. F. S. FT-NIR and linear discriminant analysis to classify chickpea seeds produced with harvest aid chemicals. *Food Chemistry*, v. 342, e128324, 2021. DOI: 10.1016/j.foodchem.2020.128324.

SACHDEV, S.; ANSARI, S. A.; ANSARI, M. I.; FUJITA, M.; HASANUZZAMAN, M. Abiotic stress and reactive oxygen species: generation, signaling, and defense mechanisms. *Antioxidants*, v. 10, e277, 2021.

SAI, V. U.; SHAMITA, M.; SREE, B. A.; APOORVA, A.; DIVYA, T.; SAI LAKSHMI, T. N. S. S.; BALRAJ, CH.; SRILATHA, A.; ARUNA KUMARI, J. Biochemical changes during seed deterioration: a review. *International Journal of Advanced Biochemistry Research*, v. 9, n. 9S, p. 1935-1944, 2025. DOI: 10.33545/26174693.2025.v9.i9Sy.5791.

SAMAMAD, N. T. I. et al. Near infrared spectroscopy, a suitable tool for fast phenotyping – the case of cashew genetic improvement. *Scientia Horticulturae*, v. 238, p. 363-368, 2018. DOI: 10.1016/j.scienta.2018.05.007.

SANTANA, F. et al. Experimento didático de quimiometria para classificação de óleos vegetais comestíveis por espectroscopia no infravermelho médio combinado com análise discriminante por mínimos quadrados parciais: um tutorial, parte v. *Química Nova*, v. 43, n. 3, p. 371-381, 2020. DOI: 10.21577/0100-4042.20170480.

SAVITZKY, A.; GOLAY, M. J. E. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, v. 36, n. 8, p. 1627-1639, 1964.

SCHUTZ, D.; RIEDL, J.; ACHTEN, E.; FISCHER, M. Fourier-transform near-infrared spectroscopy as a fast-screening tool for the verification of the geographical origin of grain maize (*Zea mays* L.). *Food Control*, v. 136, p. 108892, 2022. DOI: 10.1016/j.foodcont.2022.108892.

SENDIN, K.; MANLEY, M.; WILLIAMS, P. J. Classification of white maize defects with multispectral imaging. *Food Chemistry*, v. 243, p. 311–318, 2018.

SHAHEIN, A. M. E.; SHALABY, N. E.; MAHMOUD, B. A. Effect of storage period and condition on cotton seed viability and its chemical composition. *Journal of the Advances in Agricultural Researches*, v. 27, n. 3, p. 582-591, 2022. DOI: 10.21608/jalexu.2022.161397.1080.

SHARMA, S. N. et al. Gene expression patterns regulating the seed metabolism in relation to deterioration/ageing of primed mung bean (*Vigna radiata* L.) seeds. *Plant Physiology and Biochemistry*, v. 124, p. 40-49, mar. 2018.

SHRESTHA, S. et al. Use of multispectral imaging in varietal identification of tomato. *Sensors*, v. 15, p. 4496–4512, 2015.

SILVA, M. F.; ROQUE, J. V.; SOARES, J. M.; MOURA, L. O.; MEDEIROS, A. D.; SILVA, F. L.; SILVA, L. J. DA. Near infrared spectroscopy for the classification of vigor level of soybean seed. *Revista Ciência Agronômica*, v. 55, e20238703, 2024. DOI: 10.5935/1806-6690.20240005.

SINGH, P.; NAYYAR, A.; SINGH, S.; KAUR, A. Classification of wheat seeds using image processing and fuzzy clustered random forest. *International Journal of Agricultural Resources, Governance and Ecology*, v. 16, n. 2, p. 123, 2020. DOI: 10.1504/IJARGE.2020.109048.

SOARES, J. M.; BATISTA, T. B.; SILVA, M. F. DA; RODRIGUES, N. S.; DIAS, D. C. F. DOS S.; SILVA, L. J. DA. Classification of the physiological potential of soybean seed lots using infrared spectroscopy and chemometric methods. *Journal of Seed Science*, v. 46, 2024.

SOARES, S. F. C.; MEDEIROS, E. P.; PASQUINI, C.; MORELLO, C. DE L.; GALVÃO, R. K. H.; ARAÚJO, M. C. U. Classification of individual cotton seeds with respect to variety using near-infrared hyperspectral imaging. *Analytical Methods*, v. 8, p. 8498-8505, 2016. DOI: 10.1039/C6AY02896A.

SOUZA, L. R.; LIMÃO, M. A. R.; PINHEIRO, D. T.; PERIS, G. C. O.; DIAS, D. C. F. F.; DIAS, L. A. D. Near infrared spectroscopy and seedling image analysis to evaluate the physiological potential of *Urochloa decumbens* (Stapf) R.D. Webster seeds. *Journal of Seed Science*, v. 45, e202277021, 2023. DOI: 10.1590/2317-1545v45277021.

VIEIRA, R. D.; MARCOS-FILHO, J. Teste de condutividade elétrica. In: Krzyzanowski, F. C.; Vieira, R. D.; França-Neto, J. B.; Marcos-Filho, J. (Eds.). *Vigor de sementes: conceitos e testes*. Londrina: ABRATES, 2020, p. 333-389.

VRESAK, M. et al. The use of image-spectroscopy technology as a diagnostic method for seed health testing and variety identification. *PLoS One*, v. 11, e0152011, 2016. DOI: 10.1371/journal.pone.0152011.

WANG, D.; RAM, M. S.; DOWELL, F. E. Classification of damaged soybean seeds using near-infrared spectroscopy. *Transactions of the ASAE*, v. 45, n. 6, 2002.

WANG, L. et al. Comparative proteomics analysis reveals the mechanism of pre-harvest seed deterioration of soybean under high temperature and humidity stress. *Journal of Proteomics*, v. 75, n. 7, p. 2109-2127, abr. 2012.

WANG, M. et al. Enhancing cold and drought tolerance in cotton: a protective role of *SikCOR413PM1*. *BMC Plant Biology*, v. 23, p. 371, 2023. DOI: 10.1186/s12870-023-04572-6.

WANG, Y.; PENG, Y.; ZHUANG, Q.; ZHAO, X. Feasibility analysis of NIR for detecting sweet corn seeds vigor. *Journal of Cereal Science*, v. 93, n. 1, p. e102977, 2020.

WORKMAN Jr, J.; WEYER, L. *Practical guide to interpretive near-infrared spectroscopy*. CRC Press, 2007.

XIA, Y.; XU, Y.; LI, J.; ZHANG, C.; FAN, S. Recent advances in emerging techniques for non-destructive detection of seed viability: a review. *Artificial Intelligence in Agriculture*, v. 1, p. 35–47, 2019.

XING, Weigeng; LI, Yi; ZHOU, Linyan; HONG, Hao; LIU, Yuan; LUO, Shuailong; ZOU, Jialong; ZHAO, Yan; YANG, Yanfei; XU, Zhenjiang; TAN, Bin. Deciphering seed deterioration: molecular insights and priming strategies for revitalizing aged seeds. *Plants*, Basel, v. 14, n. 11, p. 1730, 2025. DOI: <https://doi.org/10.3390/plants14111730>.

- XU, R.; HU, W.; ZHOU, Y.; ZHANG, X.; XU, S.; GUO, Q. Use of near-infrared spectroscopy for the rapid evaluation of soybean [*Glycine max* (L.) Merri.] water soluble protein content. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, v. 224, p. 117400, 2020.
- YASMIN, J.; RAJU AHMED, M.; LOHUMI, S.; WAKHOLI, C.; KIM, M.; CHO, B.-K. Classification Method for Viability Screening of Naturally Aged Watermelon Seeds Using FT-NIR Spectroscopy. *Sensors*, v. 19, n. 5, p.e 1190, 2019.
- ZHANG, Q.; PENG, J.; WANG, J. Protective enzyme activity regulation in cotton (*Gossypium hirsutum*) under *Scirpus planiculmis* stress. *Frontiers in Plant Science*, v. 13, p. 1068419, 2022. DOI: 10.3389/fpls.2022.1068419.
- ZHENG, Q.; LIU, Q.; WANG, Y.; LAN, X.; ZHAO, Y.; LIU, P.; ZHANG, K. ABA inhibits rice seed aging by reducing H₂O₂ accumulation in the radicle. *Plants*, v. 13, n. 6, p. 809, 2024. DOI: 10.3390/plants13060809.
- ZHOU, W.; CHEN, H.; HE, L.; LU, H. Changes in antioxidant enzyme activity and lipid peroxidation in soybean seeds during ageing. *Seed Science and Technology*, v. 47, n. 1, p. 49-59, 2019. DOI: 10.15258/sst.2019.47.1.05.
- ZHUANG, T. et al. Determination of protein and fatty acid composition of shell-intact upland cottonseed using near-infrared reflectance spectroscopy. *Industrial Crops and Products*, v. 191, p. 115909, 2023. DOI: 10.1016/j.indcrop.2022.115909.