

Diferenciação de variedades de macieira por espectroscopia hiperespectral foliar e modelagem supervisionada

Differentiation of apple varieties using leaf hyperspectral spectroscopy and supervised modeling

Diferenciación de variedades de manzana mediante hiperespectroscopia foliar y modelado supervisado

Elisângela Benedet da Silva

Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina. Centro de Informações de Recursos Ambientais e de Hidrometeorologia, Florianópolis, Santa Catarina, Brasil.

 0000-0001-6288-3273

Gabriel Berenhauser Leite

Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina. Centro de Informações de Recursos Ambientais e de Hidrometeorologia, Florianópolis, Santa Catarina, Brasil.

 0000-0002-7479-7409

Marcus Vinicius Kvitschal

Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina. Estação Experimental da Epagri de Caçador, Caçador, Santa Catarina, Brasil.

 0000-0001-6161-3546

Valeria Pohlmann

Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina. Centro de Informações de Recursos Ambientais e de Hidrometeorologia, Florianópolis, Santa Catarina, Brasil.

 0000-0003-3619-599X

Cristina Pandolfo

Empresa de Pesquisa Agropecuária e Extensão Rural de Santa Catarina. Centro de Informações de Recursos Ambientais e de Hidrometeorologia, Florianópolis, Santa Catarina, Brasil.

 0000-0001-8086-7939

Resumo: A rastreabilidade varietal na cultura da macieira é estratégica para certificação de origem e proteção de cultivares, porém ainda depende de métodos laboratoriais onerosos e destrutivos. O objetivo do estudo foi determinar a viabilidade da espectroscopia hiperespectral foliar, integrada a modelos supervisionados de aprendizado de máquina, como ferramenta complementar à identificação varietal de macieira. Foram analisadas 540 folhas coletadas em pomar homogêneo de seis variedades de macieira (Fuji, Gala, Gala Gui, Galídia, Luiza e Isadora), totalizando 1620 leituras espectrais obtidas com FieldSpec® 4 e Plant Probe. Após filtragem de regiões com baixa relação sinal-ruído, aplicaram-se pré-processamentos espectrais, redução dimensional por ACP e classificação por Support Vector Classification (kernel RBF), com validação

cruzada estratificada e validação independente (hold-out, 20%). A PERMANOVA indicou efeito significativo da variedade na assinatura espectral global ($p = 0,001$). O melhor pipeline (CR + DER + MSC + ACP + SVC) apresentou f1-macro médio de 0,74 na validação cruzada. No conjunto independente, a acurácia global foi 0,67. Fuji apresentou 100% de acerto e Luiza 83%, enquanto clones do grupo Gala mostraram maior confusão (44–56%), refletindo elevada proximidade genética. A importância espectral concentrou-se nas regiões VIS, red-edge e NIR inicial, associadas a pigmentos foliares e estrutura do mesófilo, com baixa contribuição do SWIR. Os resultados demonstraram que diferenças ópticas foliares permitem discriminar cultivares, sobretudo com genealogias distintas, confirmando o potencial da abordagem como prova de conceito. Entretanto, a acurácia ainda é limitada para a certificação e rastreabilidade comercial, especialmente na separação de clones e mutantes somáticos próximos.

Palavras-chave: variedades de maçã; espectroscopia hiperespectral; reflectância foliar; aprendizado de máquina; rastreabilidade varietal.

Abstract: *Varietal traceability in apple production is essential for cultivar certification, origin authentication, and the protection of breeding programs, but it still relies on destructive and costly laboratory-based methods. This study aimed to determine the feasibility of leaf hyperspectral reflectance spectroscopy, integrated with supervised machine learning models, as a complementary tool for apple cultivar identification. A total of 540 leaves were collected from a homogeneous orchard containing six commercial apple cultivars (Fuji, Gala, Gala Gui, Galídia, Luiza, and Isadora), generating 1,620 spectral measurements acquired using a FieldSpec® 4 spectroradiometer coupled with a Plant Probe. After removing spectral regions with a low signal-to-noise ratio, spectral preprocessing, principal component analysis (PCA), and Support Vector Classification (SVC, RBF kernel) were applied using stratified cross-validation and an independent hold-out validation set (20%). PERMANOVA revealed a significant effect of cultivar on the overall spectral signature ($p = 0.001$). The best-performing pipeline (CR + DER + MSC + PCA + SVC) achieved a mean macro-F1 score of 0.74 during cross-validation. Independent validation resulted in an overall accuracy of 0.67. Fuji achieved 100% classification accuracy, whereas Luiza reached 83%, while cultivars belonging to the Gala lineage showed higher levels of confusion (44–56%), reflecting their close genetic relationship. Spectral importance was concentrated in the VIS, red-edge, and early NIR regions, which are associated with leaf pigments and mesophyll structure, whereas SWIR contributed little to cultivar discrimination. The results demonstrate that leaf optical traits can discriminate between apple cultivars, particularly those with distinct genealogies, supporting the proposed approach as a proof of concept. However, the achieved accuracy remains insufficient for commercial certification and traceability, especially for distinguishing closely related clones and somatic mutants.*

Keywords: apple varieties; hyperspectral spectroscopy; leaf reflectance; machine learning; varietal traceability.

Resumen: *La trazabilidad varietal del manzano es estratégica para la certificación de origen y la protección de cultivares, pero aún depende de métodos de laboratorio costosos y destructivos. El objetivo de este estudio fue determinar la viabilidad de la espectroscopía hiperespectral foliar, integrada con modelos supervisados de aprendizaje automático, como herramienta complementaria para la identificación varietal. Se analizaron 540 hojas recolectadas en un huerto homogéneo de seis cultivares de manzano (Fuji, Gala, Gala Gui, Galídia, Luiza e Isadora), totalizando 1.620 mediciones espectrales obtenidas con FieldSpec® 4 y Plant Probe. Tras eliminar las regiones con baja relación señal-ruido, se aplicaron preprocesamientos espectrales, análisis de componentes principales (ACP) y clasificación mediante Support Vector Classification (kernel RBF), con validación cruzada estratificada y validación independiente (hold-out, 20%). La PERMANOVA indicó un efecto significativo del cultivar sobre la firma espectral global ($p = 0,001$). El mejor procedimiento (CR + DER + MSC + ACP + SVC) obtuvo un F1-macro medio de 0,74 en la validación cruzada. En el conjunto independiente, la exactitud global fue de 0,67. Fuji alcanzó un 100% de aciertos y Luiza un 83%, mientras que los cultivares del linaje Gala presentaron mayor confusión (44–56%), debido a su proximidad genética. La importancia espectral se concentró en las regiones VIS, red-edge y NIR inicial, asociadas con los pigmentos foliares y la estructura del mesófilo, con baja contribución del SWIR. Los resultados demostraron que las propiedades ópticas foliares permiten discriminar cultivares, especialmente aquellos con genealogías distintas, confirmando RAC - Agropecuária Catarinense. 2026,*

el potencial del enfoque como prueba de concepto. Sin embargo, la exactitud aún es insuficiente para la certificación y trazabilidad comercial, particularmente para diferenciar clones y mutantes somáticos estrechamente relacionados.

Palabras clave: Cultivares de manzano; Espectroscopia hiperespectral; Reflectancia foliar; Aprendizaje automático; Trazabilidad varietal.

1 INTRODUÇÃO

A maçã (*Malus domestica* Borkh.) é uma das frutas mais cultivadas e consumidas no mundo, com destaque para a importância na alimentação humana, na indústria de alimentos e bebidas e na economia agrícola. Além de sua importância econômica, a cultura destaca-se pelo expressivo patrimônio genético disponível para programas de melhoramento (Li, 2024).

No Brasil, a produção ainda é dominada pelos grupos 'Gala' e 'Fuji', mas observa-se uma crescente incorporação de novos genótipos pela cadeia produtiva. Atualmente, Santa Catarina responde por 49% da produção brasileira (Souza; Brighenti; Goulart Junior, 2024) e mais de 60% das variedades recomendadas no Estado são de origem nacional (Faoro, 2025), evidenciando a consolidação de uma base genética cada vez mais diversificada. Nesse contexto, a identificação varietal torna-se um aspecto estratégico para a cadeia produtiva, contribuindo para a rastreabilidade, certificação de origem e proteção de novas variedades desenvolvidas pelos programas de melhoramento genético (Faoro, 2025).

A identificação varietal ainda baseia-se predominantemente em métodos tradicionais de caracterização fenotípica (descritores morfológicos e agrônômicos), cuja aplicação depende do estágio fenológico, das condições ambientais e da experiência do avaliador. Tecnologias de marcadores moleculares, tais como as análises de sequência de DNA ou RNA, os microssatélites (*Simple Sequence Repeat* - SSR) e os de polimorfismos de nucleotídeos únicos (*Simple Nucleotide Polymorphism* - SNPs), também são utilizadas e apresentam elevada confiabilidade, porém demandam infraestrutura laboratorial especializada, maior tempo de processamento e custos relativamente elevados (Sevindik *et al.*, 2024; Li, 2024; Liu *et al.*, 2022).

No escopo da agricultura de precisão, sensores hiperespectrais surgem como uma alternativa rápida e não destrutiva e potencialmente aplicável em campo, capaz de captar assinaturas espectrais associadas às propriedades biofísicas e bioquímicas dos tecidos vegetais. A diferenciação espectral baseia-se na premissa de que a variabilidade genética se traduz em assinaturas ópticas distintas permitindo o agrupamento de indivíduos com características semelhantes (Rocchini *et al.*, 2022), principalmente em função das variações na composição de pigmentos, estrutura foliar e o teor de água que afetam os padrões de reflectância, transmitância e absorbância (Carter; Knapp, 2001; Hennessy; Clarke; Lewis, 2020; Falcioni *et al.*, 2022).

Estudos recentes demonstram que assinaturas espectrais podem discriminar não apenas espécies, mas também genótipos geneticamente próximos. Diferenças espectrais entre cultivares de alface (Falcioni *et al.*, 2023), variedades de chá (Zhang *et al.*, 2020) e arroz (Yamashita *et al.*, 2020, 2021) evidenciam que variações genéticas intraespecíficas promovem alterações na composição de pigmentos foliares, na estrutura anatômica das folhas e no conteúdo de água, resultando em respostas espectrais distintas, principalmente nas regiões do visível (VIS, 400–700 nm), red-edge (680–750 nm) e infravermelho próximo (NIR, 750–1100 nm). Hennessy, Clarke e Lewis (2020) observaram que essas regiões concentram, de forma consistente, o maior poder discriminatório em estudos com vegetação, embora a seleção dos comprimentos de onda mais informativos dependa do algoritmo de classificação empregado.

No contexto das fruteiras de clima temperado, as técnicas hiperespectrais vêm se consolidando como ferramentas rápidas, não destrutivas e de elevada precisão para caracterização varietal e fenotípica. Em videira (*Vitis vinifera*), estudos utilizando espectroscopia proximal e dados hiperespectrais foliares demonstraram elevada capacidade de discriminação varietal por meio de algoritmos de aprendizado de máquina, com acurácias entre 90% e 100% (Arruda *et al.*, 2023; Mirzaei *et al.*, 2019). Esses trabalhos evidenciaram que diferenças genéticas entre cultivares resultam em alterações na composição bioquímica, na estrutura interna das folhas e em sua assinatura espectral, permitindo a identificação varietal mesmo sob condições de campo. Além disso, verificou-se que as regiões do NIR, SWIR e red-edge concentram grande parte da informação discriminatória, embora a importância relativa dos comprimentos de onda dependa da espécie e da metodologia empregada.

Em macieira, estudos utilizando espectroscopia VIS-NIR e imagem hiperespectral demonstraram elevado desempenho na identificação de cultivares, determinação da origem geográfica e avaliação de atributos de qualidade dos frutos, frequentemente alcançando acurácias superiores a 98% por meio de algoritmos de aprendizado de máquina (Cruz *et al.*, 2024; Huang *et al.*, 2020). Esses resultados evidenciam o potencial da integração entre sensores hiperespectrais e métodos quimiométricos para aplicações em autenticação varietal, rastreabilidade, fenotipagem e apoio aos programas de melhoramento genético, reduzindo a dependência de análises destrutivas e de técnicas laboratoriais baseadas em marcadores moleculares (Cruz *et al.*, 2024; Arruda *et al.*, 2023; Huang *et al.*, 2020; Mirzaei *et al.*, 2019).

Apesar dos avanços, a maior parte dos estudos em macieira concentra-se na caracterização de frutos, enquanto aplicações voltadas à discriminação de variedades por meio da reflectância foliar permanecem escassas. Essa lacuna é ainda mais evidente para cultivares adaptadas às condições edafoclimáticas do Sul do Brasil e para materiais genéticos desenvolvidos em programas nacionais de melhoramento, reforçando a necessidade de pesquisas que avaliem o potencial da espectroscopia hiperespectral foliar como ferramenta para identificação varietal em condições de campo.

Diante dessa lacuna, partiu-se da hipótese de que as diferenças genéticas entre variedades comerciais de macieira produzem padrões espectrais reprodutíveis em folhas frescas, passíveis de serem detectados por espectroscopia de reflectância e algoritmos de aprendizado de máquina. Assim, o objetivo deste estudo foi determinar a viabilidade da espectroscopia hiperespectral foliar, integrada a modelos supervisionados de aprendizado de máquina, como ferramenta complementar aos métodos convencionais de identificação varietal, por meio da avaliação da acurácia da classificação de variedades de macieira a partir de folhas provenientes de pomares localizados no estado de Santa Catarina, Sul do Brasil.

2 MATERIAIS E MÉTODOS

2.1 Área de estudo

O estudo foi conduzido em um pomar de variedades de macieira implantado em 2012, em uma área homogênea quanto ao solo, declividade e drenagem. As plantas foram conduzidas em sistema de condução do tipo líder central, com espaçamento 4×1m, sob manejo fitossanitário e nutricional padronizado. Avaliaram-se seis variedades comerciais de macieira: Fuji, Gala, Gala Gui, Galídia, Luiza e Isadora. Para cada variedade, foram amostradas cinco plantas em sequência na respectiva fileira, buscando-se manter a maior homogeneidade possível quanto à idade das plantas e às condições de manejo, solo, declividade e drenagem do pomar. Cada planta foi subdividida em três estratos verticais: inferior (0–1,2m), médio (1,2–2,0m) e superior (> 2,0m). De cada estrato, coletaram-se seis folhas maduras e saudáveis, totalizando 540 folhas. As folhas foram obtidas no período da manhã, em dias de céu limpo, nas faces de exposição norte (três folhas/estrato) e sul (três folhas/estrato) das plantas para padronizar a irradiância. As folhas foram acondicionadas em sacos de papel dentro de caixas térmicas e encaminhadas ao laboratório, sendo armazenadas em geladeira aproximadamente 15 minutos após as coletas.

2.2 Obtenção de dados espectrais

As medições espectrais foram realizadas em laboratório, sob controle de temperatura ambiente (20 e 22°C) e bloqueio da luz externa, utilizando o FieldSpec® 4 (350–2500 nm; Analytical Spectral Devices Inc., Boulder, CO, USA) acoplado ao acessório Plant Probe. O Plant Probe é equipado com fonte interna de luz halógena (12–18 V DC, 6,5 W), com temperatura de cor aproximada de 2900 K e área de leitura de 10mm de diâmetro, em geometria de contato. O sistema fornece iluminação de banda larga adequada à faixa espectral de 350–2500 nm, apresentando intensidade luminosa controlada que minimiza o aquecimento do tecido foliar e evita alterações espectrais decorrentes de estresse térmico durante as leituras (Analytical Spectral Devices Inc, 2007). Para cada folha, obtiveram-se três leituras espectrais a partir da superfície adaxial em diferentes posições da folha, totalizando 1620 leituras espectrais brutas (sem pré-processamento). As medições foram corrigidas com o painel de referência Spectralon® no início e a cada 20 minutos de leitura. Todas as leituras espectrais foram exportadas e organizadas utilizando o software *ViewSpec Pro* (Analytical Spectral Devices Inc.,

Boulder, CO, USA). As três leituras realizadas em cada folha foram agregadas pela média, acompanhadas do desvio-padrão ($\pm 1\sigma$), resultando em uma única curva espectral representativa por folha.

2.3 Filtragem espectral e análise PERMANOVA

As curvas espectrais originais (350–2500 nm), compostas por 2151 bandas, foram submetidas à remoção de 565 bandas localizadas nas regiões de baixa relação sinal-ruído (SNR; <400 nm e >2350 nm) e nas principais bandas de absorção da água ($\sim 970 \pm 30$ nm, $\sim 1400 \pm 50$ nm e $\sim 1850 \pm 100$ nm). Como resultado, 1586 bandas de maior integridade espectral foram utilizadas nas análises posteriores (Hennessy; Clarke; Lewis, 2020). Em seguida, calculou-se a matriz de distâncias Euclidianas entre as 540 folhas, considerando simultaneamente todas as bandas filtradas. A matriz de distâncias foi submetida à análise PERMANOVA, com 999 permutações, para testar se os centroides espectrais difeririam entre as seis variedades de macieira. A PERMANOVA permitiu avaliar a separação multivariada entre as variedades no espaço espectral, sem realizar testes univariados banda a banda.

2.4 Pré-processamento dos dados espectrais

Quatro pré-processamentos espectrais foram avaliados isoladamente ou combinados entre si. Os pré-processamentos *Multiplicative Scatter Correction* (MSC), *Standard Normal Variate* (SNV), *Continuum Removal* (CR) e *Derivada de Savitzky–Golay* (DER) foram aplicados nas curvas espectrais brutas visando corrigir efeitos de dispersão óptica, padronizar variações estruturais das folhas e realçar feições espectrais associadas aos pigmentos e à estrutura celular, conforme descrito na literatura sobre tratamento de espectros NIR/hiperespectrais (Rinnan; Van Den Berg; Engelsen, 2009; Zhang *et al.*, 2022; Li *et al.*, 2025).

2.5 Análise de Componentes Principais e seleção de componentes

As combinações possíveis (MSC, SNV, CR e DER) foram avaliadas sistematicamente por permutações, incluindo também as curvas espectrais brutas (sem pré-processamento). Para cada pré-processamento, aplicou-se uma análise de componentes principais (ACP) com retenção de 90%, 95% ou 99% da variância total. Os escores da ACP foram submetidos a uma ANOVA entre as variedades de macieira para identificar componentes principais com diferenças significativas ($p < 0.001$), sendo apenas essas componentes utilizadas nos modelos subsequentes. A ANOVA foi recalculada internamente em cada *fold* da validação cruzada, garantindo que apenas os dados do conjunto de treinamento fossem utilizados na seleção das componentes principais e evitando vazamento de informação (*data leakage*). Como a ACP é uma transformação não supervisionada, seus *loadings* permaneceram estáveis entre os *folds*, não incorporando informação das classes. Quando nenhum componente significativo foi detectado, todas as PCs retidas pela variância foram mantidas. Os escores selecionados foram posteriormente padronizados (*StandardScaler*) antes da classificação.

2.6 Classificação e validação

A classificação das variedades de macieira foi realizada com o algoritmo variando o coeficiente C (0,1; 1; 10) e gamma (“scale” e “auto”). Para a classificação das variedades de macieira, utilizou-se o algoritmo *Support Vector Classification* (SVC, *kernel* RBF), uma implementação do *Support Vector Machine* (SVM) amplamente reconhecido na literatura para classificação de dados hiperespectrais. Segundo Grewal *et al.* (2023), o SVM destaca-se pela sua capacidade de processar todos com alta dimensionalidade, como as bandas espectrais e de modelar relações não-lineares entre assinaturas espectrais foliares, sendo indicado para cenários onde a diferenciação entre as classes é sutil e o número de amostras de treino é limitado, garantindo uma melhor generalização do modelo. Para avaliar o desempenho das diferentes combinações de pré-processamento, redução dimensional e classificação, construíram-se *pipelines* (fluxos integrados de processamento e modelagem), nos quais cada etapa de pré-processamento, redução dimensional e classificação foram executadas de forma sequencial.

A avaliação inicial dos *pipelines* foi conduzida por validação cruzada estratificada ($k = 5$), calculando acurácia e f1-macro. Os dez *pipelines* com maior f1-macro foram registrados. O modelo final (*pipeline* de melhor desempenho) foi avaliado particionando de forma estratificada o conjunto de dados para preservar a proporção original de folhas por variedade de macieira (90 folhas por cultivar). O conjunto de treinamento com 432 folhas (80% dos dados) e de validação independente (*hold-out*, 20% dos dados), com 108 folhas não utilizadas na etapa de treinamento. Assim, cada classe contribuiu com 18 folhas no conjunto de validação. O desempenho preditivo foi quantificado por meio de métricas de precisão, *recall* e *f1-score* por variedade de macieira e pela matriz de confusão. O *recall* (sensibilidade) indica a capacidade do modelo em identificar corretamente amostras pertencentes a cada variedade de macieira. A matriz de confusão permitiu visualizar, em termos absolutos, como cada variedade foi classificada e identificada, os padrões de acerto e confusão entre classes.

2.7 Importância espectral

A contribuição relativa de cada banda espectral para o desempenho do classificador foi estimada por *Permutation Importance* (30 permutações) utilizando o conjunto independente de validação. Esse método quantificou a queda na métrica f1-macro ao embaralhar individualmente cada banda, identificando regiões espectrais sensíveis ao processo de diferenciação varietal. Os picos de importância foram detectados automaticamente a partir da distribuição das importâncias ($>90^{\text{a}}$ percentil). Todas as análises foram realizadas em Python v3.12.13, utilizando principalmente os pacotes numpy (v2.0.2), pandas (v2.2.2), scikit-learn, scipy (v1.16.3), matplotlib (v3.10.0) e seaborn (v0.13.2) (Python Software Foundation, 2026).

De forma geral, o fluxo de análise consistiu em: organização das curvas espectrais das folhas, remoção das regiões com maior ruído, aplicação de diferentes tratamentos nos espectros, redução da dimensionalidade por ACP, seleção dos componentes principais mais informativos e treinamento do modelo de classificação. Em seguida, o desempenho do modelo foi avaliado com folhas não utilizadas no treinamento, permitindo estimar a capacidade real de distinguir as variedades de macieira.

Ferramentas de inteligência artificial baseadas em modelos de linguagem, incluindo o *ChatGPT* (OpenAI) e o *SciSpace*, foram utilizadas como apoio à redação do manuscrito, organização textual, revisão gramatical e identificação preliminar de literatura científica. A busca e seleção das referências foram complementadas por pesquisas no Google Acadêmico, e todas as referências citadas foram verificadas pelos autores em suas fontes originais. A geração, o processamento e a interpretação dos dados, bem como as análises estatísticas e as conclusões, são de responsabilidade exclusiva dos autores.

3 RESULTADOS E DISCUSSÕES

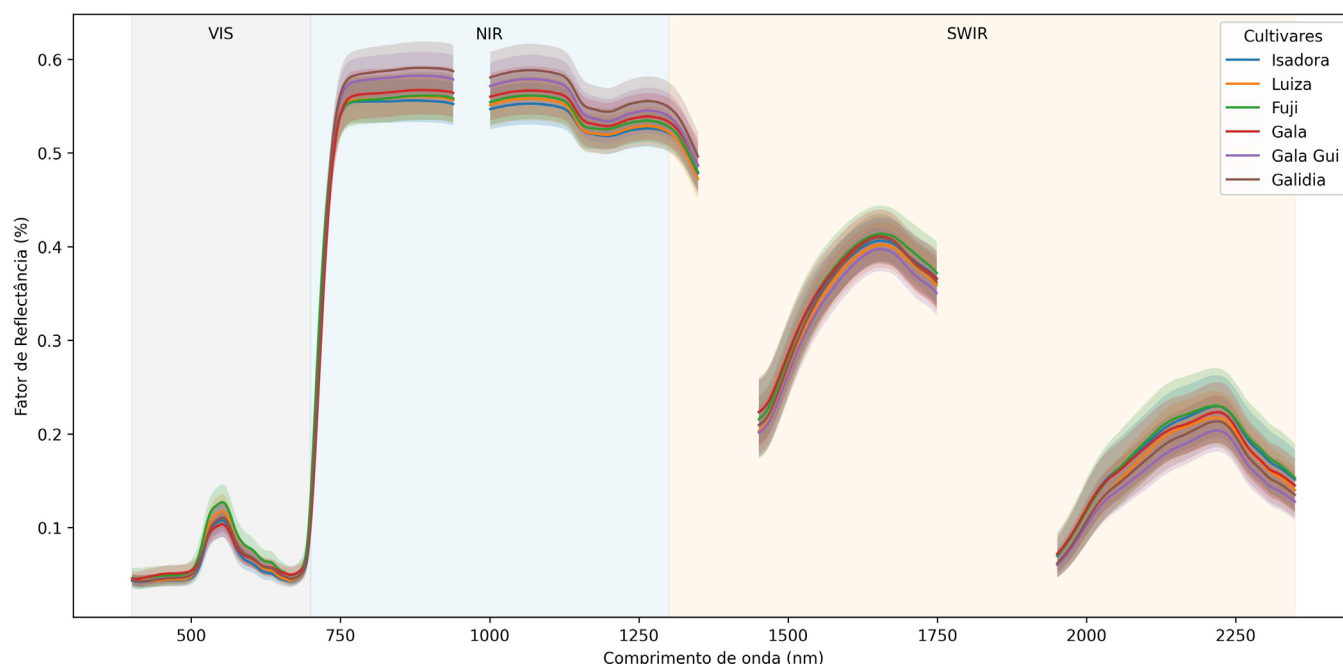
Os resultados indicam que a espectroscopia hiperespectral foliar, combinada à modelagem supervisionada, é capaz de discriminar cultivares de macieira com desempenho elevado quando há maior divergência genética e/ou funcional entre os materiais, enquanto a separação de clones mutantes se mostra mais desafiadora devido à proximidade genômica e à consequente similaridade de traços opticamente ativos. Nessas situações, diferentes características biológicas podem ocupar regiões muito próximas no espaço espectral, aproximando-se do conceito de “tipos ópticos” discutido na literatura de diversidade espectral (Rocchini *et al.*, 2022). A interpretabilidade do modelo pode ser apresentada pela identificação de comprimentos de onda mais relevantes (VIS, *red-edge* e NIR) e pela caracterização dos principais pares de confusão, apoiando regras operacionais baseadas em probabilidade de classe e categorias “incertas”.

Na Figura 1 estão apresentadas as curvas médias ($\pm 1\sigma$ desvio-padrão) de reflectância espectral foliar das diferentes variedades de macieira ao longo das regiões do VIS, infravermelho próximo (NIR) e infravermelho de ondas curtas (SWIR). Os dados evidenciam padrões espectrais clássicos de folhas verdes, com baixa reflectância no VIS, transição abrupta no *red edge*, alta reflectância no NIR e feições de absorção bem definidas no SWIR. Embora as curvas médias apresentem elevada sobreposição visual ao longo do espectro, observam-se variações sutis e sistemáticas entre

as variedades. Essas diferenças foram avaliadas por meio de PERMANOVA aplicada à matriz de distâncias espectrais, que indicou efeito significativo da variedade sobre a assinatura espectral global (pseudo-F = 16,54; p = 0,001; 999 permutações). Esse resultado demonstra que, quando consideradas simultaneamente em um espaço espectral multidimensional, as variações entre variedades superam significativamente a variabilidade intravarietal.

Assim, ainda que as diferenças banda a banda sejam discretas, sua integração ao longo de todo o espectro resulta em separação multivariada estatisticamente significativa entre as variedades. Em nível foliar, essa abordagem preserva assinaturas espectrais “puras”, permitindo maximizar o potencial discriminante associado aos traços bioquímicos e estruturais das plantas (Blackburn, 2007; Hennessy; Clarke; Lewis, 2020; Zhang *et al.*, 2022, Rocchini *et al.*, 2022).

Figura 1 - Curvas espectrais médias de reflectância foliar (%) (linhas) dos variedades de macieira Isadora, Luiza, Fuji, Gala, Gala Gui e Galidia, com desvio-padrão ($\pm 1\sigma$, áreas sombreadas), ao longo do espectro eletromagnético (400–2400nm). As faixas coloridas indicam as regiões do VIS, NIR e SWIR. As lacunas espectrais correspondem aos intervalos removidos devido à baixa relação sinal-ruído (SNR), incluindo <400nm, >2350nm e regiões de absorção da água (~970, ~1400 e ~1850nm)



Fonte: elaborado pelos autores (2025)

Na região do VIS, observa-se baixa reflectância, geralmente inferior a 10%, com exceção do pico do verde, em torno de 550 nm, seguindo padrão típico de folhas saudáveis. As diferenças entre as curvas são mais evidentes nas regiões do azul (400–450 nm) e do verde (500–550 nm), onde a dispersão entre variedades, embora pequena em valor absoluto, é consistente e superior ao desvio-padrão intravarietal. Essa região é fortemente controlada pela absorção dos pigmentos fotossintéticos, com predominância da clorofila a e b, seguida pelos carotenoides e antocianinas, cujos efeitos espectrais tendem a ser parcialmente mascarados pela clorofila (Blackburn, 2007; Hennessy; Clarke; Lewis, 2020). A transição abrupta entre o mínimo de reflectância no vermelho (680 nm) e o platô do NIR evidencia claramente o red edge em todas as variedades de macieira. A estabilidade da curva espectral na posição do red edge indica ausência de estresse severo ou senescência avançada na folha, sendo as variações observadas atribuídas majoritariamente a diferenças varietais intrínsecas (Su, 2020; Zhang *et al.*, 2022), e não respostas fisiológicas transitórias.

No NIR, todas as variedades de macieira apresentam elevada reflectância (55–60%), característica da forte dispersão interna da radiação pelo mesófilo foliar. Essas variações estão associadas a diferenças na espessura da folha, organização celular e proporção de espaços intercelulares, conforme descrito nos modelos de transferência radiativa foliar (Zhang *et al.*, 2022).

No SWIR, as diferenças espectrais entre variedades são discretas e amplamente sobrepostas pelo desvio-padrão, o que explica a baixa importância do SWIR na classificação varietal entre plantas de mesma espécie. Esse comportamento está de acordo com Hennessy, Clarke e Lewis (2020), que apontam maior relevância do SWIR em estudos voltados à composição química e ao teor de água, e não à discriminação varietal em folhas fisiologicamente semelhantes. Na Tabela 1 são apresentados os resultados das 10 melhores pipelines. O f1-macro foi empregado como métrica principal na etapa de validação cruzada estratificada, para a comparação global dos pipelines por refletir o desempenho equilibrado entre as classes.

Tabela 1 – Desempenho dos dez melhores pipelines de pré-processamento espectral e classificação, avaliados por validação cruzada estratificada ($k = 5$). Os pipelines diferem quanto às combinações de pré-processamentos, parâmetros do filtro de Savitzky-Golay e número de componentes principais (PCs) retidas para explicar 99% da variância

<i>Pipeline</i>	<i>wl</i>	<i>po</i>	<i>n_pcs</i>	<i>SVC</i> (<i>C, gamma</i>)	<i>Acurácia</i> <i>média</i>	<i>f1-macro (%)</i>
CR + DER + MSC	11	2	12	1.0, scale	0,7403 ± 0,04	0,7365 ± 0,04
CR + DER + SNV	11	2	12	1.0, scale	0,7366 ± 0,04	0,7331 ± 0,04
CR + DER	11	2	10	10.0, auto	0,7255 ± 0,05	0,7240 ± 0,05
CR + SNV + DER	15	3	12	1.0, scale	0,7255 ± 0,07	0,7213 ± 0,07
CR + MSC	NaN	NaN	8	1.0, scale	0,7217 ± 0,05	0,7188 ± 0,04
CR + SNV + MSC	NaN	NaN	8	1.0, auto	0,7217 ± 0,05	0,7188 ± 0,04
CR + SNV + DER	7	2	11	1.0, auto	0,7217 ± 0,02	0,7162 ± 0,02
CR + MSC + DER	15	3	12	10.0, auto	0,7144 ± 0,05	0,7125 ± 0,05
CR + MSC + DER	15	3	12	1.00, scale	0,7163 ± 0,08	0,7124 ± 0,08
CR + SNV + DER	15	3	12	10.0, auto	0,7125 ± 0,05	0,7106 ± 0,05

Fonte: elaborado pelos autores (2025).

Nota: # *wl*: comprimento da janela (número de bandas) do filtro de Savitzky-Golay; *po*: ordem do polinômio do filtro de Savitzky-Golay; *n_pcs*: número de componentes principais retidas (99% da variância explicada); *SVC* (*C, gamma*): parâmetros do classificador Support Vector Classification com kernel RBF, onde *C* é o coeficiente de penalização e *gamma* define o alcance da influência de cada ponto de treinamento; *Acurácia média*: proporção média de classificações corretas; *f1-macro*: média não ponderada do f1-score calculado individualmente para cada classe, atribuindo igual peso a todas as variedades.

Entre os pré-processamentos avaliados, as combinações CR, DER (window=11, polyorder=2) + MSC (acurácia = 0.740; f1-macro = 0.736) e CR, DER (window=11, polyorder=2) +e SNV (acurácia = 0.736; f1-macro = 0.733) apresentaram os melhores desempenhos de classificação, com diferenças discretas (< 0,5 ponto percentual) entre as métricas obtidas. O CR destaca feições pigmentares e estruturais, a derivada enfatiza mudanças espectrais relacionadas à anatomia foliar e tanto o MSC quanto o SNV atuam na correção final de espalhamento e *drift* instrumental, amplamente empregados em espectroscopia de reflectância difusa (Savitzky; Golay, 1964; Pasquini, 2003; Rinnan; Van Den Berg; Engelsen, 2009; Li *et al.*, 2025). A ACP, adotando limiar de 99% de variância explicada, resultou na retenção de 16 componentes principais (PCs), dentre um total máximo possível de 538 eixos. Esse número corresponde ao conjunto mínimo de componentes necessários para atingir o critério de variância acumulada estabelecido.

A aplicação subsequente de análise de variância univariada (ANOVA, teste F), com critério de significância ($p < 0,001$), indicou que 12 desses componentes apresentaram relevância estatística para a discriminação varietal, sendo mantidos para a etapa de classificação. Esse resultado sugere que a variabilidade espectral capturada pela ACP está, ao menos em parte, associada às diferenças entre as variedades de macieira. A retenção apenas dos componentes estatisticamente significativos indica que os eixos utilizados na modelagem apresentam contribuição discriminativa relevante, embora não se possa afirmar que estejam completamente isentos de variabilidade residual ou redundante, mas sim que demonstraram pertinência estatística no contexto da classificação proposta. Assim, o modelo final foi construído com um subconjunto de componentes estatisticamente informativos, promovendo uma representação parcimoniosa da variabilidade espectral disponível. No espaço reduzido, o classificador SVC, utilizando hiperparâmetros padrão ($C = 1$; $\gamma = \text{scale}$), apresentou desempenho consistente, alcançando f1-macro médio de 0,718 em validação cruzada estratificada ($k = 5$).

O uso de algoritmos de aprendizado de máquina para o processamento dos dados hiperespectrais justifica-se pela natureza complexa e não-linear das assinaturas ópticas das variedades de macieira. Conforme revisado por Grewal *et al.* (2023), técnicas de aprendizado de máquina são fundamentais para lidar com a alta dimensionalidade dos dados hiperespectrais e extrair padrões não lineares associados às assinaturas espectrais varietais.

A Figura 2 mostra a importância espectral relativa das variáveis obtidas por *permutation importance* a partir do conjunto de validação (*hold-out*, 20%). Nesse contexto, a importância espectral quantifica a redução no desempenho do modelo quando os valores de uma determinada banda são aleatoriamente permutados no conjunto de validação. Essa permutação rompe a relação entre a banda e a variável resposta (variedade), de modo que uma queda acentuada no desempenho indica maior relevância preditiva daquela região espectral para a discriminação varietal.

As bandas mais informativas concentraram-se na região do VIS (400-700 nm), com os maiores picos na região em torno do azul/azul-borda do verde (400-560 nm), com o pico do verde centrado em torno de 550 nm e a região do vermelho tardio/transição para o red edge em torno dos 680 nm. Nessa região, os pigmentos clorofila a e b, carotenoides e antocianinas, presentes nas folhas, dominam a reflectância espectral e o percentual desses pigmentos pode variar bastante. Essas variações indicam diferenças na absorção desses pigmentos e são amplamente exploradas por classificadores espectrais na discriminação entre variedades (Hennessy; Clarke; Lewis, 2020).

Os picos no red edge, início do NIR, em torno de 680–696 nm, correspondem à inflexão característica associada às variações de clorofila, mas também tem sido associada ao conteúdo de antocianina, além de estresse e estrutura foliar, enquanto que em torno de 700 nm aparece um pico relevante, refletindo sensibilidade à espessura do mesófilo foliar (Falcioni *et al.*, 2023). No NIR, as bandas em torno de 938 nm e 1003 nm apontam características de espalhamento celular (estrutura interna da folha) que também contribuíram para a diferenciação das variedades de macieira.

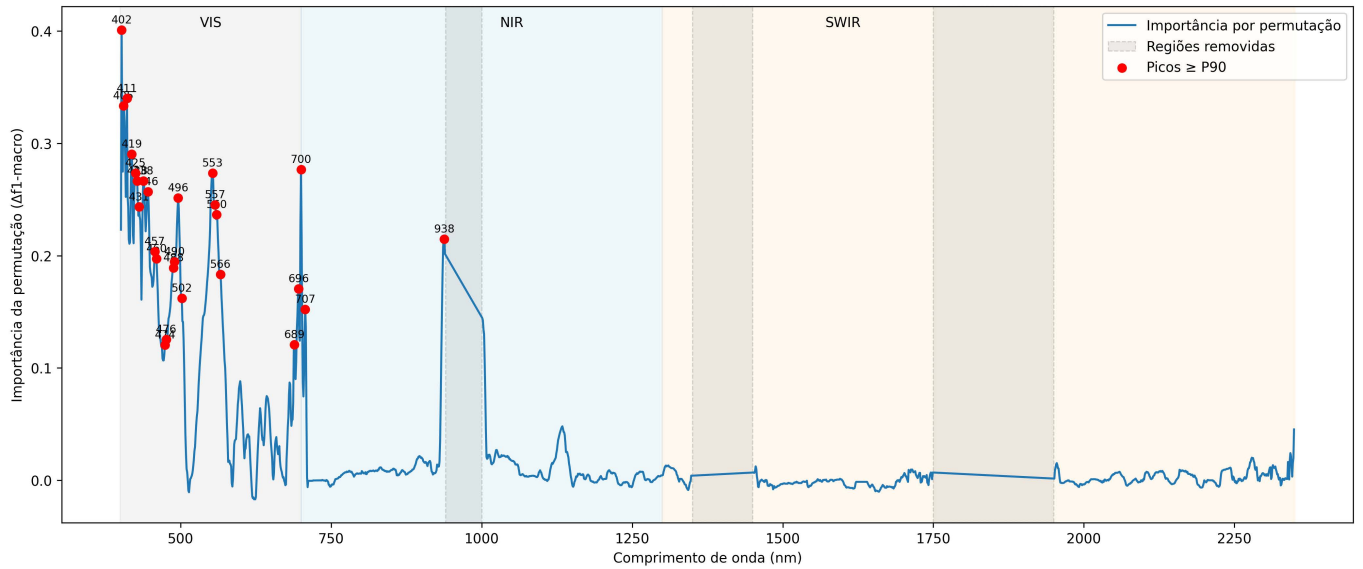
Nas regiões acima de 1100 nm, a importância para a classificação varietal cai drasticamente, aproximando-se de zero, o que justifica a remoção dessas bandas em muitos modelos devido ao baixo ganho de sinal informativo. Como as folhas analisadas eram fisiologicamente semelhantes e não apresentavam sinais de estresse severo (estabilidade do red edge, Figura 1), o conteúdo de água e a biomassa seca que regem a reflectância no SWIR apresentaram-se quase uniformes entre as diferentes variedades de macieira, reduzindo o poder discriminativo desta região. Esse comportamento confirma que, em condições fisiológicas semelhantes às regiões dominadas por pigmentos no VIS e pela estrutura celular no NIR concentram maior poder discriminativo do que as bandas do SWIR, cuja variabilidade é majoritariamente governada por vibrações O–H associadas à água (Zhang *et al.*, 2020; Yamashita *et al.*, 2021; Falcioni *et al.*, 2022).

O desempenho do modelo de classificação foi avaliado por meio da matriz de confusão (Figura 3) e das métricas de precisão, recall e f1-score, calculadas a partir de um conjunto de validação independente (*hold-out*, 20% dos dados - 18 folhas por variedade). O recall foi interpretado como a sensibilidade do modelo em identificar corretamente folhas pertencentes a cada variedade de macieira, enquanto o f1-score representa os respectivos desempenhos específicos no modelo final. Adicionalmente, foram consideradas a acurácia global (accuracy), representando a proporção total de classificações corretas, e o f1-score médio por macro-average, que atribui peso igual a todas as variedades de macieira, tendo sido consideradas seis delas com mesmo número de amostras por classe.

No conjunto de validação, o modelo atingiu acurácia global de 0,67, com f1-score médio de 0,67, indicando desempenho equilibrado entre as seis classes avaliadas. Embora moderada, essa acurácia é comparável à de estudos recentes que diferenciam variedades vegetais de mesma espécie por espectroscopia (Falcioni *et al.*, 2022; Cushnahan *et al.*, 2024).

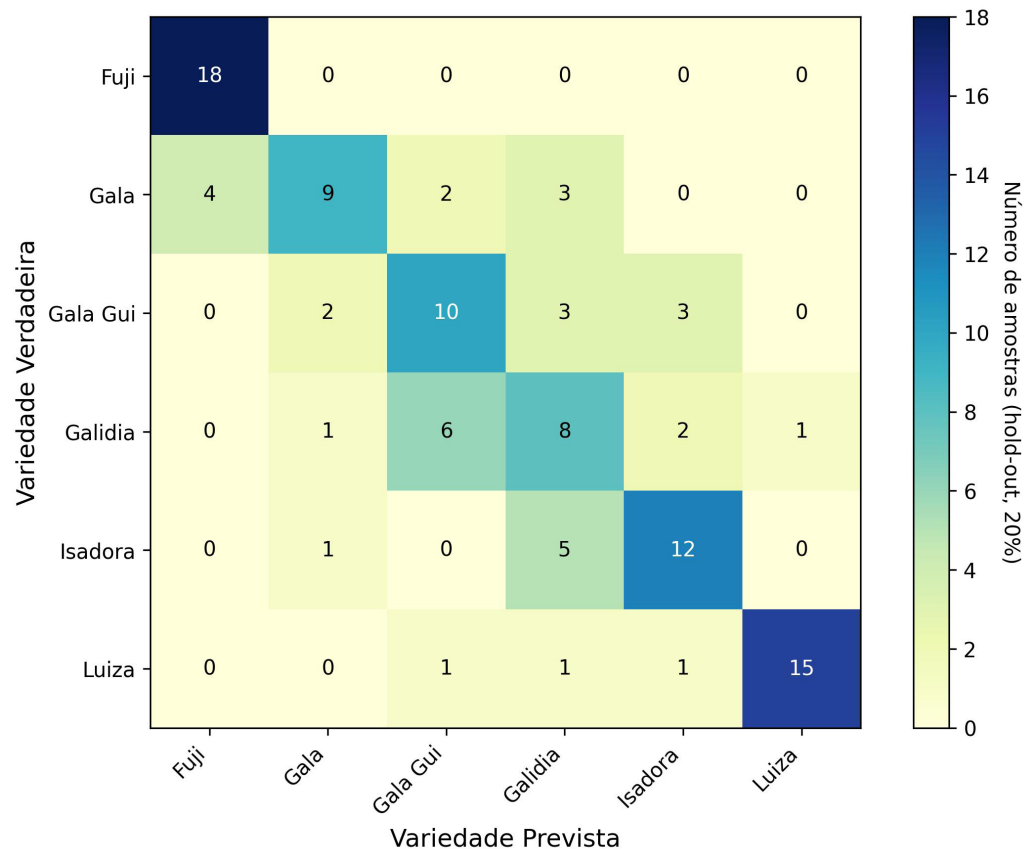
A matriz de confusão mostrou diferenças no desempenho de classificação entre as variedades de macieira objeto deste estudo (Figura 3). A variedade Fuji apresentou 100% de acerto no conjunto de validação (18/18 amostras), indicando elevada distinção espectral em relação às demais. A Gala apresentou desempenho intermediário (9/18), com confusões principalmente com Fuji (4 amostras) e com os clones Gala Gui (2 amostras) e Galídia (3 amostras), refletindo similaridades espectrais esperadas entre genótipos. A variedade Luiza alcançou 15/18 acertos, com confusões pontuais e equilibradas com Isadora, Galídia e Gala Gui (1 amostra cada), indicando boa separabilidade global. Isadora apresentou 12/18 acertos, sendo a maior parte das confusões associada à Galídia (5 amostras), além de uma confusão com Gala. As maiores taxas de erro foram observadas entre Gala Gui e Galídia, ambos clones da variedade Gala. Gala Gui apresentou

Figura 2 - Importância espectral determinada por *permutation importance* no conjunto de teste (hold-out, 20%) para o pipeline otimizado, ao longo do espectro (400–2350nm). Os pontos vermelhos marcam os picos de importância (> 90% do valor máximo), destacando comprimentos de onda na porção VIS (400–700nm), red-edge (680–780nm) e NIR (900–1100nm). As lacunas espectrais correspondem aos intervalos removidos devido à baixa relação sinal-ruído (SNR), incluindo <400nm, >2350nm e regiões de absorção da água (~970, ~1400 e ~1850nm)



Fonte: elaborado pelos autores (2025)

Figura 3 - Matriz de confusão do classificador aplicada ao conjunto de validação independente (hold-out, 20% do conjunto de dados), ilustrando o desempenho na discriminação das variedades de macieira Fuji, Gala, Gala Gui, Galidia, Isadora e Luiza. As linhas representam as variedades verdadeiras e as colunas as variedades previstas pelo modelo (SVC). Os valores indicam o número de amostras por classe, enquanto a intensidade da coloração reflete a frequência de acertos e erros de classificação



Fonte: elaborado pelos autores (2025)

10/18 acertos, com confusões recorrentes principalmente com Galídia e Isadora (3 amostras cada), enquanto Galídia apresentou 8/18 acertos, sendo fortemente confundida com Gala Gui (6 amostras).

Esse padrão reforça a elevada similaridade espectral entre materiais geneticamente muito próximos, onde as diferenças fisiológicas e estruturais foliares são extremamente sutis. Rocchini *et al.* (2022), exemplificam ainda que mesmo indivíduos de genótipos distintos podem compartilhar propriedades bioquímicas similares, resultando em uma sobreposição severa no ‘espaço de traços funcionais’. Nesses casos, a semelhança dos traços opticamente ativos supera a capacidade de diferenciação taxonômica do sensor, agrupando diferentes entidades biológicas em um mesmo ‘tipo óptico’ ou ‘espécie espectral’. As variedades Fuji e Luiza apresentaram os melhores resultados, com valores elevados de *recall* (1,00 e 0,83, respectivamente) e *f1-score* superiores a 0,88, indicando assinaturas espectrais bem definidas e elevada consistência intravarietal.

Esses resultados sugerem que essas variedades apresentam padrões pigmentares e estruturais mais distintos no espaço espectral analisado. Por outro lado, as variedades Gala, Gala Gui e Galídia apresentaram desempenho intermediário inferior, com *recall* variando entre 0,44 e 0,56 e *f1-scores* entre 0,42 e 0,58. Os erros de classificação concentram-se principalmente entre essas variedades, indicando confusão espectral parcial, especialmente entre Gala Gui e Galídia, e entre Gala, Gala Gui e Galídia. Esse padrão é coerente com a elevada proximidade genética e morfológica dessas variedades, bem como com a sobreposição observada nas curvas espectrais médias, sobretudo nas regiões do VIS e do *red edge*. Esse elevado grau de confundimento entre essas variedades de macieira já era esperado, uma vez que ambas são clones mutantes da variedade Star Gala que é um clone mutante de Royal Gala que, por sua vez, é um clone mutante da variedade Gala original, conforme árvores genealógicas descritas por Faoro (2025) e por Denardi *et al.* (2023). As variedades Gala Gui e Galídia são clones mutantes de Gala, e são geneticamente muito similares entre si, apresentando apenas variações nos genes relacionados ao padrão de cobertura de coloração vermelha na epiderme das frutas e resistência a doenças, sem variações visuais nos padrões pigmentares das folhas.

A variedade Isadora apresentou desempenho intermediário (*recall* e *f1-score* = 0,67), com erros de classificação predominantemente associados à Galídia, sugerindo similaridade espectral localizada, possivelmente relacionada a características comuns de pigmentação foliar e estrutura do mesófilo. A matriz de confusão também indicou erros majoritariamente restritos a grupos varietais próximos, não havendo confusões generalizadas entre todas as classes. Isso reforça que o modelo capturou padrões espectral e biologicamente consistentes, mesmo quando a separação entre variedades foi sutil. Novamente ressalta-se a coerência dos resultados, sugerindo que isso se deve à proximidade genealógica entre as variedades. Variedades com ancestrais comuns, particularmente aquelas derivadas da linhagem Gala e com recorrente contribuição de Golden Delicious (Isadora, Galídia e Gala Gui) apresentaram maior sobreposição espectral e, conseqüentemente, maiores taxas de confusão entre si. Em contraste, variedades como Fuji e Luiza exibiram assinaturas espectrais mais distintas, refletindo-se em maiores valores de precisão e *recall*.

No caso da variedade Luiza, embora esta seja neta de Fuji e detenha aproximadamente 25% de seu genoma, a macieira é uma espécie alógama e altamente heterozigota, na qual a segregação genética ao longo das gerações promove amplo embaralhamento alélico. Esse processo resulta em combinações genéticas únicas em cada descendente, podendo levar à expressão fenotípica distinta mesmo entre variedades com ancestralidade compartilhada. Considerando que os padrões espectrais foliares são determinados por múltiplos caracteres quantitativos, incluindo pigmentação, estrutura do mesófilo e teor hídrico, controlados por numerosos loci com possíveis efeitos aditivos, dominantes ou epistáticos, é esperado que a segregação em duas gerações de cruzamentos produza assinaturas espectrais diferenciadas. Assim, a baixa taxa de confusão entre Luiza e Fuji não decorre de menor compartilhamento genético, mas da recombinação e segregação poligênica inerentes ao melhoramento da macieira.

Os confundimentos observados entre a variedade Gala e seus clones (Galídia e Gala Gui) são consistentes com a estreita proximidade genética entre esses materiais. Os cultivares Galídia e Gala Gui são mutações somáticas espontâneas derivadas de Star Gala, pertencendo à mesma linhagem genética iniciada pelo cultivar Gala (Faoro, 2025; Kvitschal *et al.*, 2023). Nesses casos, o elevado compartilhamento do fundo genético tende a resultar em elevada similaridade estrutural e bioquímica das folhas, produzindo assinaturas espectrais muito próximas e limitando a capacidade discriminativa do sensor. Por outro lado, a confusão observada entre variedades com genealogias distintas, como Isadora e Galídia

(Figura 3), sugere que a separação espectral não depende exclusivamente do grau de parentesco entre os materiais, mas também da expressão de atributos funcionais foliares capazes de gerar respostas ópticas semelhantes. Essa interpretação é compatível com a hipótese proposta por Rocchini *et al.* (2022) de que genótipos distintos podem compartilhar propriedades ópticas e funcionais semelhantes. Entretanto, esse mecanismo não foi avaliado diretamente no presente estudo e deverá ser investigado em pesquisas futuras. Informações detalhadas sobre a genealogia, ancestralidade comum e grau de parentesco entre os cultivares avaliados encontram-se descritas em Faoro (2025) e Kvitschal *et al.* (2023).

A elevada acurácia observada para cultivares de genealogias distintas (f1-score > 85%) reforça que diferenças genéticas podem modular propriedades ópticas detectáveis por espectroscopia (Yamashita *et al.*, 2021; Falcioni *et al.*, 2023), enquanto a elevada similaridade genética entre clones e mutantes somáticos, aliada a pequenas variações fisiológicas e fenológicas, provavelmente limitou a discriminação entre materiais estreitamente aparentados. A hipótese de que a resposta espectral foliar pode distinguir variedades de macieira foi parcialmente confirmada. Embora a acurácia obtida no conjunto independente ainda não seja suficiente para aplicações operacionais de certificação e rastreabilidade comercial, para as quais são desejáveis níveis mais elevados de confiabilidade, não há atualmente um patamar regulatório estabelecido para validação de métodos hiperespectrais aplicados à identificação varietal de macieira. Assim, os resultados devem ser interpretados como uma prova de conceito para o desenvolvimento de uma ferramenta complementar de triagem rápida e não destrutiva, ampliando o potencial de análises em larga escala sem substituir os métodos moleculares de referência.

Quanto à aplicação operacional de sensores multiespectrais e hiperespectrais em condições de campo, especialmente aqueles embarcados em drones, a extrapolação das assinaturas espectrais obtidas em folhas para o nível de copa é tecnicamente factível, porém requer etapas adicionais de calibração radiométrica, segmentação das copas, controle do estágio fenológico e adaptação de domínio para minimizar os efeitos da mistura de alvos, da iluminação natural e da geometria de aquisição, conforme amplamente reconhecido na literatura de sensoriamento remoto aplicado a pomares (Zhang *et al.*, 2021).

4 CONCLUSÃO

A espectroscopia hiperespectral foliar, integrada a modelos supervisionados de aprendizado de máquina, demonstrou viabilidade para a discriminação varietal de macieira, confirmando que diferenças ópticas associadas às folhas permitem distinguir cultivares, sobretudo aqueles com genealogias distintas. As regiões espectrais do VIS, red-edge e início do NIR concentraram o maior poder discriminatório, reforçando sua relevância para estudos de classificação varietal.

A abordagem mostrou potencial como ferramenta complementar de triagem rápida e não destrutiva aos métodos convencionais de identificação varietal. Entretanto, a acurácia obtida no conjunto independente ainda não é suficiente para aplicações operacionais de certificação e rastreabilidade comercial, especialmente na diferenciação de clones e mutantes somáticos, cuja elevada similaridade genética resultou em assinaturas espectrais muito próximas. Assim, os resultados devem ser interpretados como uma prova de conceito para o desenvolvimento de métodos espectrais aplicados à identificação varietal.

O estudo também demonstrou que a resposta espectral foliar não depende exclusivamente do grau de parentesco entre os cultivares, indicando que fatores fisiológicos e funcionais podem contribuir para a sobreposição de assinaturas espectrais entre alguns materiais. Esses mecanismos, entretanto, não foram avaliados diretamente e permanecem como uma importante lacuna para futuras investigações.

Estudos futuros deverão avaliar a estabilidade temporal das assinaturas espectrais ao longo do ciclo fenológico, sob diferentes condições ambientais e hídricas, bem como ampliar o conjunto de cultivares e validar a metodologia em diferentes pomares. A integração da espectroscopia hiperespectral com métodos moleculares de referência poderá ampliar a robustez da identificação varietal, conciliando escalabilidade, rapidez e confiabilidade em programas de rastreabilidade e certificação.

REFERÊNCIAS

ANALYTICAL SPECTRAL DEVICES INC. **ASD Accessories User Manual**. Boulder, CO: ASD Inc., 2007.

ARRUDA, D. C.; DUCATI, J. R.; HOFF, R.; BELLOLI, T. F.; THUM, A. B. Proximal hyperspectral analysis in grape leaves for region and variety identification. **Ciência Rural**, v. 53, n. 12, e20220313, 2023. DOI: <http://doi.org/10.1590/0103-8478cr20220313>.

BLACKBURN, G. A. Hyperspectral remote sensing of plant pigments. **Journal of Experimental Botany**, Oxford, v. 58, n. 4, p. 855–867, 2007. DOI: <http://doi.org/10.1093/jxb/erl123>.

CARTER, G. A.; KNAPP, A. K. Leaf optical properties in higher plants: linking spectral characteristics to stress and chlorophyll concentration. **American Journal of Botany**, v. 88, n. 4, p. 677–684, 2001. DOI: <https://doi.org/10.2307/2657068>.

CRUZ, J.; GERGOV, G.; TARAPOULOUZI, M.; KIRILOVA, E.; KOSTADINOVA, O.; NIKOLOVA, K. Classification of apple varieties by VIS-NIR diffuse reflectance spectroscopy and chemometrics. **Bulgarian Chemical Communications**, v. 56, p. 106-115, 2024. DOI: <https://doi.org/10.34049/bcc.56.D.S1P38>.

CUSHNAHAN, T. A.; GRAFTON, M. C. E.; PEARSON, D.; RAMILAN, T. Hyperspectral data can differentiate species and cultivars of C3 and C4 Turf despite measurable diurnal variation. **Remote Sensing**, v. 16, p. 3142, 2024. DOI: <https://doi.org/10.3390/rs16173142>.

DENARDI, F.; KVITSCHAL, M. V.; HAWERROTH, M. C.; PETRI, J. L. **Cultivares de macieira desenvolvidos pela Epagri**. Florianópolis: Epagri, 2023. 97 p. (Boletim Técnico, 211).

FALCIONI, R.; ANTUNES, W. C.; DEMATTÊ, J. A. M.; NANNI, M. R. Reflectance spectroscopy for the classification and prediction of pigments in agronomic crops. **Plants**, v. 12, p. 2347, 2023. DOI: <https://doi.org/10.3390/plants12122347>.

FALCIONI, R.; GONÇALVES, J. V. F.; OLIVEIRA, K. M. D.; ANTUNES, W. C.; NANNI, M. R. VIS-NIR-SWIR hyperspectroscopy combined with data mining and machine learning for classification of predicted chemometrics of green lettuce. **Remote Sensing**, v. 14, p. 6330, 2022. DOI: <https://doi.org/10.3390/rs14246330>.

FAORO, I. D. (org.). **Características agronômicas de cultivares de macieira**. Florianópolis, SC: Epagri, 2025. 211 p. Disponível em: <https://publicacoes.epagri.sc.gov.br/livros/article/view/2200>. Acesso em: 24 out. 2025.

GREWAL, R.; KASANA, S. S.; KASANA, G. Machine learning and deep learning techniques for spectral spatial classification of hyperspectral images: A comprehensive survey. **Electronics**, v. 12, p. 488, 2023. DOI: <https://doi.org/10.3390/electronics12030488>.

HENNESSY, A.; CLARKE, K.; LEWIS, M. Hyperspectral classification of plants: A review of waveband selection generalisability. **Remote Sens.**, v. 12, p. 113, 2020. DOI: <https://doi.org/10.3390/rs12010113>.

HUANG, Y.; YANG, Y.; SUN, Y.; ZHOU, H.; CHEN, K. Identification of Apple Varieties Using a Multichannel Hyperspectral Imaging System. **Sensors**, v. 20, n. 18, p. 5120, 2020. DOI: <https://doi.org/10.3390/s20185120>.

KVITSCHAL, M. V.; DENARDI, F.; HAWERROTH, M. C.; PETRI, J. L. **Cultivares de macieira desenvolvidos pela Epagri**. Florianópolis, 2023. 97 p. (Boletim Técnico, 211).

LI, J. Genetic diversity and conservation strategies of apple germplasm resources. **Tree Genetics and Molecular Breeding**, v. 14, n. 1, p. 1-7, 2024. DOI: <https://doi.org/10.5376/tgmb.2024.14.0001>.

LI, Y.; XU, X.; WU, W.; ZHU, Y.; GAO, L.; JIANG, X.; MENG, Y.; Yang, G.J.; XUE, H. Hyperspectral estimation of chlorophyll content in grapevine based on feature selection and GA-BP. **Scientific Reports**, v. 15, p. 8029, 2025. DOI: <https://doi.org/10.1038/s41598-024-84977-x>.

LIU, W.; CHEN, Z.; JIANG, S.; WANG, Y.; FANG, H.; ZHANG, Z.; CHEN, X.; WANG, N. Research Progress on Genetic Basis of Fruit Quality Traits in Apple (*Malus × domestica*). **Frontiers in Plant Science**, v. 13, p. 918202, 2022. DOI: <https://doi.org/10.3389/fpls.2022.918202>.

MIRZAEI, M.; MAROFI, S.; ABBASI, M.; SOLGI, E.; KARIMI, R.; VERRELST, J. Scenario-based discrimination of common grapevine varieties using in-field hyperspectral data in the western of Iran. **International Journal of Applied Earth Observation and Geoinformation**, v. 80, p. 26-37, 2019. DOI: <https://doi.org/10.1016/j.jag.2019.04.002>.

PASQUINI, C. Near infrared spectroscopy: fundamentals, practical aspects and analytical applications. **Journal of the Brazilian Chemical Society**, São Paulo, v. 14, p. 198-219, 2003. DOI: <https://doi.org/10.1590/S0103-50532003000200006>.

RINNAN, Å.; VAN DEN BERG, F.; ENGELSEN, S. B. Review of the most common pre-processing techniques for near-infrared spectra. **Trends in Analytical Chemistry**, v. 28, n. 10, p. 1201-1222, 2009. DOI: <https://doi.org/10.1016/j.trac.2009.07.007>.

ROCCHINI, D.; SANTOS, M. J.; USTIN, S. L.; FÉRET, J.-B.; ASNER, G. P.; BEIERKUHNLEIN, C.; DALPONTE, M.; FEILHAUER, H.; FOODY, G. M.; GELLER, G. N.; GILLESPIE, T. W.; HE, K. S.; KLEIJN, D.; LEITÃO, P. J.; MALAVASI, M.; MOUDRÝ, V.; MÜLLEROVÁ, J.; NAGENDRA, H.; NORMAND, S.; RICOTTA, C.; SCHAEPMAN, M. E.; SCHMIDTLEIN, S.; SKIDMORE, A. K.; ŠÍMOVÁ, P.; TORRESANI, M.; TOWNSEND, P. A.; TURNER, W.; VIHervaara, P.; WEGMANN, M.; LENOIR, J. The spectral species concept in living color. **Journal of Geophysical Research: Biogeosciences**, v. 127, e2022JG007026, 2022. DOI: <https://doi.org/10.1029/2022JG007026>.

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. DOI: <https://doi.org/10.1021/ac60214a047>.

SEVINDIK, E.; YANGILAR, F.; ATASAGUN, B.; SOFYALIOĞLU, E.; ŞİMŞEK, S.A.; ÇAYIR, M.E.; VIVODÍK, M. Genetic diversity and structure analysis of 'Ak Saki' and 'Kara Saki' apple cultivars growing in Erzincan/Türkiye. **Biologica Nyssana**, v. 15, p. 61-68, 2024. DOI: <https://doi.org/10.5281/zenodo.14258237>.

SOUZA, R. P.; BRIGHENTI, A. F.; GOULART JUNIOR, R. Análise da evolução da produção de maçã em Santa Catarina no período de 2014 a 2021. **Agropecuária Catarinense**, Florianópolis, v. 37, n. 3, p. 40-45, 2024. DOI: <https://doi.org/10.52945/rac.v37i3.1801>.

SU, W.-H. Advanced Machine Learning in Point Spectroscopy, RGB- and Hyperspectral-Imaging for Automatic Discriminations of Crops and Weeds: A Review. **Smart cities**, v. 3, p. 767-792, 2020. DOI: <http://doi.org/10.3390/smartcities3030039>. Acesso em: 01 abr. 2025.

YAMASHITA, H.; SONOBE, R.; HIRONO, Y.; MORITA, A.; IKKA, T. Dissection of hyperspectral reflectance to estimate nitrogen and chlorophyll contents in tea leaves based on machine learning algorithms. **Scientific Reports**, v. 10, p. 17360, 2020. DOI: <https://doi.org/10.1038/s41598-020-73745-2>.

YAMASHITA, H.; SONOBE, R.; HIRONO, Y.; MORITA, A.; IKKA, T. Potential of spectroscopic analyses for non-destructive estimation of tea quality-related metabolites in fresh new leaves. **Scientific Reports**, v. 11, p. 4169, 2021. DOI: <https://doi.org/10.1038/s41598-021-83847-0>.

ZHANG, C.; VALENTE, J.; KOOISTRA, L.; GUO, L.; WANG, W. Orchard management with small unmanned aerial vehicles: a survey of sensing and analysis approaches. **Precision Agriculture**, v. 22, p. 2007–2052, 2021. DOI: <https://doi.org/10.1007/s11119-021-09813-y>.

ZHANG, J.; WANG, C.; YUAN, L.; LIU, P.; ZHANG, Y.; WU, K.. Construction of a plant spectral library based on an optimised feature selection method. **Biosystems Engineering**, v. 195, p. 1–16, 2020. DOI: <https://doi.org/10.1016/j.biosystemseng.2020.04.008>.

ZHANG, Y.; LI, X.; WANG, C.; ZHANG, R.; JIN, L.; HE, Z.; TIAN, S.; WU, K.; WANG, F. PROSPECT-PMP+: Simultaneous Retrievals of Chlorophyll a and b, Carotenoids and Anthocyanins in the Leaf Optical Properties Model. **Sensors**, v. 22, p. 3025, 2022. DOI: <https://doi.org/10.3390/s22083025>.

INFORMAÇÕES ADICIONAIS

Breve currículo dos autores:

Elisângela Benedet da Silva

Doutora em Ciência do Solo (UFRGS). Mestrado em Engenharia Civil (UFSC). Pesquisadora na Empresa Agropecuária Catarinense - Epagri, no Centro de Informações de Recursos Ambientais e de Hidrometeorologia. E-mail. elisangelasilva@epagri.sc.gov.br

Gabriel Berenhauser Leite

Doutor em Fisiologia Vegetal (UBP-França). Mestrado em Agronomia (UFPel). Pesquisador na Empresa Agropecuária Catarinense - Epagri, no Centro de Informações de Recursos Ambientais e de Hidrometeorologia. E-mail. gabriel@epagri.sc.gov.br

Marcus Vinicius Kvitschal

Doutor em Genética e Melhoramento (UEM). Mestrado em Agronomia (UEM). Pesquisador na Empresa Agropecuária Catarinense - Epagri, na Estação Experimental de Caçador. E-mail. marcusvinicius@epagri.sc.gov.br

Valéria Pohlmann

Doutora em Agronomia (UFPel). Mestrado em Agronomia (UFSM). Pesquisadora na Empresa Agropecuária Catarinense - Epagri, no Centro de Informações de Recursos Ambientais e de Hidrometeorologia. E-mail. valeriapohlmann@epagri.sc.gov.br

Cristina Pandolfo

Doutora em Recursos Genéticos (UFSC). Mestrado em Fitotecnia (UFRGS). Pesquisadora na Empresa Agropecuária Catarinense - Epagri, no Centro de Informações de Recursos Ambientais e de Hidrometeorologia. E-mail. cristina@epagri.sc.gov.br

Contribuição de autoria - CRediT:

Elisângela Benedet da Silva: conceituação; curadoria; análise formal; obtenção de financiamento; investigação; metodologia; recursos; software; supervisão; validação; visualização dos dados; escrita- primeira redação; escrita- revisão e edição.

Gabriel Berenhauser Leite: conceituação; investigação; escrita- primeira redação; escrita- revisão e edição.

Marcus Vinicius Kvitschal: conceituação; investigação; escrita- primeira redação; escrita- revisão e edição.

Valéria Pohlmann: investigação; escrita- primeira redação; escrita- revisão e edição.

Cristina Pandolfo: investigação; Obtenção de financiamento; escrita- revisão e edição.

Financiamento:

Este trabalho conta com financiamento do CNPq (404759/2021-1) e Epagri/Ciram.

Conflitos de interesse:

Os autores declaram não haver conflitos de interesses neste trabalho.

Licença de uso: Mesma revista

Preprints: Não

Aprovação do comitê de ética na investigação: Não se aplica

Consentimento para o uso de imagens: Sim

Disponibilidade dos dados de investigação e outros materiais:

Os dados e demais materiais de investigação estão disponíveis mediante solicitação aos autores.

Acordo com a revisão aberta: Sim

Agradecimentos:

Os autores agradecem à Estação Experimental da Epagri de Caçador pelo apoio às atividades de campo que viabilizaram a execução desta pesquisa e ao CNPq pelo financiamento (404759/2021-1).