

UNIVERSIDADE FEDERAL DE VIÇOSA

**Segmentação semântica de lavouras de café utilizando arquitetura U-Net e
imagens de satélite**

Daniel Henrique Leite
Magister Scientiae

**VIÇOSA - MINAS GERAIS
2025**

DANIEL HENRIQUE LEITE

**Segmentação semântica de lavouras de café utilizando arquitetura U-Net e
imagens de satélite**

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Engenharia Agrícola, para obtenção do título de *Magister Scientiae*.

Orientador: Domingos S. M. Valente

Coorientadores: Daniel M. de Queiroz
Diego Bedin Marin
Fábio Daniel Tancredi

**VIÇOSA - MINAS GERAIS
2025**

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

T

L533s
2025
Leite, Daniel Henrique, 1998-
Segmentação semântica de lavouras de café utilizando
arquitetura U-Net e imagens de satélite / Daniel Henrique Leite.
– Viçosa, MG, 2025.

1 dissertação eletrônica (54 f.): il. (algumas color.).

Orientador: Domingos Sárvio Magalhães Valente.
Dissertação (mestrado) - Universidade Federal de Viçosa,
Departamento de Engenharia Agrícola, 2025.

Referências bibliográficas: f. 45-54.

DOI: <https://doi.org/10.47328/ufvbbt.2026.003>

Modo de acesso: World Wide Web.

1. Café - Cultivo - Sensoriamento remoto. 2. Aprendizado
do computador. 3. Análise espectral. 4. Café - Zona da Mata
(MG : Mesorregião) . I. Valente, Domingos Sárvio Magalhães,
1978-. II. Universidade Federal de Viçosa. Departamento de
Engenharia Agrícola. Programa de Pós-Graduação em
Engenharia Agrícola. III. Título.

CDD 22. ed. 633.73580285631

DANIEL HENRIQUE LEITE

Segmentação semântica de lavouras de café utilizando arquitetura U-Net e imagens de satélite

Dissertação apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Engenharia Agrícola, para obtenção do título de *Magister Scientiae*.

APROVADA: 4 de agosto de 2025.

Assentimento:

Daniel Henrique Leite
Autor

Domingos Sarvio Magalhaes Valente
Orientador

Essa dissertação foi assinada digitalmente pelo autor em 07/01/2026 às 12:36:58 e pelo orientador em 06/02/2026 às 10:04:24. 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 **GV5W.IKKM.9DG8** e clique no botão 'Validar documento'.

Aos meus pais Agostinho e Zenite
Dedico

AGRADECIMENTOS

Em primeiro lugar, agradeço aos meus pais Agostinho e Zenite, que me ensinaram os princípios éticos para que eu me tornasse um homem digno e íntegro, por estarem sempre ao meu lado e por me apoiarem em todos os momentos da vida. Sou grato também por acreditarem que este sonho poderia se tornar realidade.

Ao meu irmão Alexandre e à sua família, que sempre me apoiaram, ajudaram e incentivaram.

Ao amor da minha vida, Tailaine, por ser minha inspiração e por todo o companheirismo, compreensão, apoio e carinho.

Aos meus amigos Rodrigo Lopes (in memoriam) e João Bosco Soares, pelo apoio e incentivo nos estudos desde o ensino médio.

Ao meu orientador, Prof. Dr. Domingos Sárvio Magalhães Valente, expresse minha sincera gratidão pelos conselhos, pela orientação e por todo o apoio desde o início desta jornada na graduação.

Aos meus coorientadores, Prof. Dr. Daniel Marçal de Queiroz, Prof. Dr. Diego Bedin Marin e ao Pesq. Dr. Fábio Daniel Tancredi, pela significativa contribuição para o desenvolvimento da minha formação científica.

À Empresa de Pesquisa Agropecuária de Minas Gerais (EPAMIG), pela importante contribuição ao desenvolvimento deste trabalho, em especial a Pedro Maya e Ryan Borges.

Por fim, gostaria de agradecer à Universidade Federal de Viçosa, em especial ao Departamento de Engenharia Agrícola, pela excelente estrutura e pela oportunidade de formação.

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).

“A persistência é o caminho do êxito”
(Charles Chaplin)

RESUMO

LEITE, Daniel Henrique, M.Sc., Universidade Federal de Viçosa, agosto de 2025. **Segmentação semântica de lavouras de café utilizando arquitetura U-Net e imagens de satélite.** Orientador: Domingos Sarvio Magalhaes Valente. Coorientadores: Daniel Marcal de Queiroz, Diego Bedin Marin e Fábio Daniel Tancredi.

A estimativa de áreas cultivadas com café é fundamental para o planejamento estratégico da cafeicultura e elaboração de políticas públicas, o que beneficia todo o setor. Este estudo teve como objetivo segmentar lavouras de café na região das Matas de Minas, estado de Minas Gerais, Brasil. Para isto, foram selecionadas sete cenas e extraídas 54 imagens do satélite PlanetScope entre os anos de 2021 e 2024, priorizando baixa cobertura de nuvens. Utilizou-se a arquitetura U-Net para segmentação semântica de uso e cobertura do solo, com foco na identificação de lavouras de café. As imagens das cenas 2 e 7 foram utilizadas para treinamento e validação, enquanto as demais serviram para teste. A rotulagem das imagens dessas duas cenas foi feita manualmente seguindo a abordagem de rotulagem densa para quatro classes: Coffee Crops, Forest, City e Other, totalizando 1.264.844.800 pixels classificados. Para as cinco cenas independentes (1, 3, 4, 5 e 6), utilizadas no conjunto de teste, aplicou-se a mesma metodologia de rotulagem densa empregada no treinamento e validação. As imagens foram divididas em patches de 256x256 pixels e submetidas a normalização. Foram testadas diferentes combinações de bandas espectrais: (B, G, R), (B, G, NIR), (B, R, NIR) e (G, R, NIR), visando entender a influência de combinações de bandas na segmentação por classe. O modelo foi treinado por 10 épocas com early stopping e redução de taxa de aprendizado, utilizando o otimizador Adam e função de perda categorical crossentropy. A estrutura da U-Net foi adaptada para entradas de 256x256 pixels e quatro classes de saída, mantendo as conexões skip e a simetria entre os caminhos codificador e decodificador. A rede U-Net apresentou desempenho consistente na segmentação das classes definidas, com destaque para a combinação espectral (B, G, R), que obteve os melhores resultados gerais no conjunto de validação com IoU de 0,9354 e F1-score de 0,9663. No entanto, a combinação (G, R, NIR) foi a que melhor identificou a classe Coffee Crops, com IoU de 0,9260 e F1-score de 0,9616 para a classe. A análise das predições para as quatro combinações espectrais evidenciou que os principais erros ocorreram entre Coffee Crops e Other, indicando sobreposição espectral entre essas classes. O modelo final, que apresentou o melhor desempenho na classe Coffee Crops, que foi com a combinação (G, R, NIR), foi aplicado nas

imagens das cenas do conjunto de teste e manteve desempenho geral consistente, com IoU de 0,8082 e F1-score de 0,8929, com resultados satisfatórios para a classe Coffee Crops, com IoU de 0,7851 e F1-score de 0,8796. As demais classes, especialmente City e Forest, apresentaram alto desempenho. Os resultados confirmam a capacidade de generalização do modelo, e também que a combinação espectral influencia diretamente na segmentação, com as combinações (B, G, R) e (G, R, NIR) destacando-se, sendo que, a presença simultânea das bandas G e R se mostraram essenciais para discriminação da classe Coffee Crops e a combinação (G, R, NIR) demonstrou ser a mais eficaz para identificar a classe. No entanto, a confusão na predição entre Coffee Crops e Other, impacta na delimitação espacial precisa das classes, resultando em valores discrepantes de F1-score e IoU. O que evidencia os desafios inerentes à segmentação exata de classes espectralmente semelhantes. Esta interpretação é crucial para compreender que o modelo mantém uma boa capacidade de identificação, mas que a precisão de seus contornos pode ser aprimorada em trabalhos futuros.

Palavras-chave: Cafeicultura, Segmentação Semântica, Sensoriamento Remoto, Agricultura Digital, Combinação de bandas.

ABSTRACT

LEITE, Daniel Henrique, M.Sc., Universidade Federal de Viçosa, August, 2025. **Semantic segmentation of coffee crops using U-Net architecture and satellite imagery.** Adviser: Domingos Sarvio Magalhaes Valente. Co-advisers: Daniel Marcal de Queiroz, Diego Bedin Marin and Fábio Daniel Tancredi.

The estimation of areas cultivated with coffee is fundamental for the strategic planning of coffee farming and the development of public policies, which benefits the entire sector. This study aimed to segment coffee crops in the Matas de Minas region, in the state of Minas Gerais, Brazil. For this purpose, seven scenes were selected and 54 images were extracted from the PlanetScope satellite between 2021 and 2024, prioritizing low cloud coverage. The U-Net architecture was used for semantic segmentation of land use and land cover, focusing on coffee crop identification. Scenes 2 and 7 were used for training and validation, while the others served for testing. The labeling of the images from these two scenes was done manually following a dense labeling approach for four classes: Coffee Crops, Forest, City, and Other, totaling 1,264,844,800 classified pixels. For the five independent scenes (1, 3, 4, 5, and 6) used in the test set, the same dense labeling methodology applied in the training and validation was used. The images were divided into 256x256 pixel patches and subjected to normalization. Different combinations of spectral bands were tested: (B, G, R), (B, G, NIR), (B, R, NIR), and (G, R, NIR), aiming to understand the influence of band combinations on class segmentation. The model was trained for 10 epochs with early stopping and learning rate reduction, using the Adam optimizer and the categorical crossentropy loss function. The U-Net structure was adapted for 256x256 pixel inputs and four output classes, maintaining the skip connections and the symmetry between the encoder and decoder paths. The U-Net network showed consistent performance in segmenting the defined classes, with the (B, G, R) spectral combination standing out, as it achieved the best overall results on the validation set with an IoU of 0.9354 and an F1-score of 0.9663. However, the (G, R, NIR) combination was the most effective at identifying the Coffee Crops class, with an IoU of 0.9260 and an F1-score of 0.9616 for that class. The analysis of the predictions for the four spectral combinations showed that the main errors occurred between Coffee Crops and Other, indicating spectral overlap between these classes. The final model, which demonstrated the best performance for the Coffee Crops class using the (G, R, NIR) combination, was applied to the images from the test set and maintained consistent overall performance, with an IoU of 0.8082 and an F1-score of 0.8929, achieving satisfactory results for

the Coffee Crops class with an IoU of 0.7851 and an F1-score of 0.8796. The other classes, especially City and Forest, showed high performance. The results confirm the model's generalization capability and also that the spectral combination directly influences segmentation, with the (B, G, R) and (G, R, NIR) combinations standing out. The simultaneous presence of the G and R bands proved to be essential for discriminating the Coffee Crops class, and the (G, R, NIR) combination was shown to be the most effective for identifying this class. However, the prediction confusion between Coffee Crops and Other impacts the precise spatial delimitation of the classes, resulting in discrepant F1-score and IoU values. This highlights the inherent challenges of accurately segmenting spectrally similar classes. This interpretation is crucial for understanding that while the model maintains a good identification capability, the precision of its boundaries can be improved in future work.

Keywords: Coffee Crops, Semantic Segmentation, Remote Sensing, Digital Agriculture, Band Combination.

LISTA DE FIGURAS

Figura 1. Localização da área de estudo. A região em azul corresponde às Matas de Minas, enquanto as áreas em vermelho destacam as sete cenas selecionadas (Cenas 1 a 7). Dessas cenas foram obtidas 54 imagens mensais do sensor PlanetScope, adquiridas entre 2021 e 2024.	23
Figura 2. Arquitetura U-Net adaptada e utilizada neste trabalho. O <i>Input</i> corresponde a patches de 256×256 pixels com três bandas espectrais, e o <i>Output</i> apresenta a máscara de segmentação com quatro classes: Coffee Crops (vermelho), Forest (verde), City (cinza) e Other (preto). ..	28
Figura 3. Matrizes de confusão geradas para cada combinação de bandas espectrais utilizadas como entrada na arquitetura U-Net, avaliadas no conjunto de validação: (a) combinação de bandas (B, G, R); (b) combinação de bandas (B, G, NIR); (c) combinação de bandas (B, R, NIR); e (d) combinação de bandas (G, R, NIR).	32
Figura 4. Resultados da segmentação de imagens do conjunto de validação utilizando as combinações espectrais (B, G, R), (B, G, NIR), (B, R, NIR) e (G, R, NIR), com base na arquitetura U-Net. As cores indicam as classes: Coffee Crops (vermelho), Forest (verde), City (cinza) e Other (preto).	34
Figura 5. Exemplos de previsões do modelo treinado aplicado nas cenas de teste (1, 3, 4, 5 e 6). As cores representam as classes previstas pelo modelo: Other (amarelo), Coffee Crops (vermelho), Forest (verde) e City (cinza).	36
Figura 6. Matriz de confusão obtida a partir da segmentação das cinco cenas de teste pelo modelo com a combinação de bandas (G, R, NIR).	38

LISTA DE TABELAS

Tabela 1. Distribuição de pixels rotulados por classe no conjunto de dados para treinamento e validação	25
Tabela 2. Resultados gerais da segmentação da combinação de diferentes bandas com base no conjunto de validação	30
Tabela 3. Avaliação da combinação de bandas por classe com base na F1-score e no conjunto de validação	31
Tabela 4. Avaliação da combinação de bandas por classe com base na IoU e no conjunto de validação	31
Tabela 5. Desempenho por classe do modelo com combinação de bandas (G, R, NIR) nas cinco cenas de teste, com valores de F1-score e IoU	37

SUMÁRIO

1	INTRODUÇÃO	13
2	REVISÃO DE LITERATURA.....	15
	2.1 CAFEICULTURA	15
	2.2 SENSORIAMENTO REMOTO.....	17
	2.3 MACHINE LEARNING	18
	2.4 SEGMENTAÇÃO SEMÂNTICA.....	19
	2.5 ESTUDOS RELACIONADOS	20
	2.5.1 Baseados em <i>Machine Learning</i>	20
	2.5.2 Baseados em Classificador da Máxima Verossimilhança	21
	2.5.3 Utilização de sensores ativos na identificação de lavouras de café....	22
3	MATERIAL E MÉTODOS	22
	3.1 ÁREA DE ESTUDO E IMAGENS UTILIZADAS	22
	3.2 DESENVOLVIMENTO DO MODELO	24
	3.2.1 Rotulagem.....	24
	3.2.2 Pré-processamento dos dados	26
	3.2.3 Modelagem	27
	3.2.4 Métricas de Avaliação	28
4	RESULTADOS	29
	4.1 MODELO DESENVOLVIDO	30
	4.2 ANÁLISE DOS PADRÕES DE PREDIÇÃO ENTRE AS CLASSES.....	31
	4.3 ANÁLISE COMPARATIVA DAS MÁSCARAS DE SEGMENTAÇÃO ...	33
	4.4 TESTE DO MODELO EM NOVAS CENAS.....	35
5	DISCUSSÃO	39
	5.1 DESEMPENHO GLOBAL DAS COMBINAÇÕES ESPECTRAIS NA SEGMENTAÇÃO	39
	5.2 INFLUÊNCIA DAS BANDAS ESPECTRAIS NA SEGMENTAÇÃO POR CLASSE	40
	5.3 INTERPRETAÇÃO DA CONFUSÃO ESPECTRAL ENTRE CLASSES...	42
6	CONCLUSÃO.....	43
	REFERÊNCIAS.....	45

1 INTRODUÇÃO

O café, além de ser uma das bebidas mais populares e consumidas diariamente no mundo, é uma das commodities agrícolas mais importantes no comércio internacional (Spence; Carvalho, 2020). A produção de café é uma atividade importante para a subsistência de aproximadamente 25 milhões de famílias em todo o mundo, sendo majoritariamente realizada por pequenos produtores que têm essa cultura como principal fonte de renda. Além disso, o café desempenha um papel econômico fundamental em diversas nações, sendo cultivado em cerca de 80 países localizados em regiões tropicais (DaMatta *et al.*, 2019).

Diante dessa relevância, a estimativa precisa da área cultivada com café torna-se fundamental. Ela orienta desde o planejamento estratégico da safra pelos produtores, como a compra antecipada de defensivos e a alocação de recursos para a área cultivada, como também é necessária para prever antecipadamente a produção de café. Além dos benefícios diretos para os produtores, essas estimativas também desempenham um papel fundamental para o governo. Elas são ferramentas essenciais para proteger a cafeicultura e o comércio nacional, permitindo a elaboração de políticas públicas que visem estabilizar o mercado interno, além de antecipar riscos associados a desastres (Kouadio *et al.*, 2021)

Tradicionalmente, o mapeamento de áreas cultivadas com café tem sido realizado de forma manual por meio de ferramentas de sistemas de informação geográfica (SIG), além do uso de classificadores supervisionados e análise estatística de imagens provenientes de sensoriamento remoto. Técnicas como o classificador da máxima verossimilhança (Martínez-Verduzco; Galeana-Pizaña; Cruz-Bello, 2012; Moreira; Barros; Rudorff, 2008), o algoritmo *Random Forest* (Bourgoin *et al.*, 2020a) e métodos baseados em textura como a Gray-Level Co-occurrence Matrix (GLCM) (Silva *et al.*, 2009a) são utilizados. Apesar de sua relevância, esses métodos enfrentam limitações em mapeamentos escaláveis, especialmente devido à alta variabilidade visual das lavouras de café e à dificuldade em capturar sua complexidade semântica.

Essa tarefa se torna ainda mais complexa devido à variação em tamanho, forma e cor das lavouras de café, o que dificulta a distinção em relação a outras classes de cobertura do solo (De Carvalho Alves *et al.*, 2023; Schmitt-Harsh, 2013a). Essa complexidade leva os algoritmos tradicionais de segmentação, como os baseados em limiarização, arestas, regiões e gráficos, a apresentarem desempenho limitado (Wang *et al.*, 2023). Além de possuírem elevado custo

computacional, esses métodos nem sempre conseguem equilibrar eficiência e precisão, mostrando-se pouco adequados para lidar com a especificidade visual das plantações de café.

Nesse contexto, o uso de *Deep Learning* tem ganhado destaque como uma solução para identificar com maior eficiência áreas de cultivos (Chang *et al.*, 2023). Entre as técnicas mais promissoras para essa aplicação está a segmentação semântica, uma abordagem que consiste na tarefa de classificar cada pixel de uma imagem atribuindo-lhe uma classe específica, o que permite a identificação precisa de diferentes classes de cobertura do solo. Entre os modelos mais utilizados nesse contexto, destaca-se a U-Net, uma *Deep Neural Network* projetada especificamente para tarefas de segmentação semântica. Sua arquitetura em formato de “U” combina etapas de *encoder*, responsáveis por extrair características progressivamente mais complexas, com etapas de *decoder*, que recuperam a resolução espacial da imagem e permitem uma classificação precisa em nível de pixel (Ronneberger; Fischer; Brox, 2015a).

Estudos recentes indicam que, especialmente em dados com alta resolução espacial e baixa qualidade espectral, as *Deep Neural Networks* se destacam por captar características semânticas mais complexas nas imagens (Ayhan *et al.*, 2020; Zhang *et al.*, 2020). Essa abordagem contribui para superar uma limitação recorrente na classificação da cobertura do solo, relacionada à ocorrência de elementos visualmente semelhantes com assinaturas espectrais distintas, e de elementos distintos com assinaturas espectrais semelhantes (Chang *et al.*, 2023). Com o uso de técnicas de *Deep Learning*, torna-se possível extrair informações mais robustas, resultando em ganhos significativos na precisão da identificação de culturas específicas, como o café.

Além da capacidade das *Deep Neural Networks* de extrair características semânticas complexas, a seleção de bandas espectrais utilizadas como entrada nos modelos é um fator determinante para a eficácia na segmentação da cobertura do solo (Bhuiyan *et al.*, 2020). Isso porque diferentes tipos de cobertura do solo, como vegetação, solos e culturas agrícolas podem apresentar comportamentos distintos de refletância ao longo das faixas espectrais, o que permite, por meio de combinações específicas de bandas, a discriminação de classes de cobertura do solo mesmo em cenários de difícil distinção (Ding *et al.*, 2023; Hatfield *et al.*, 2019; Jia *et al.*, 2024). No entanto, essas redes normalmente restringem a entrada a apenas três canais espectrais, o que torna a escolha da combinação de bandas um fator determinante, que por sua vez, influencia diretamente no resultado final da segmentação.

A seleção de bandas espectrais torna-se ainda mais relevante em regiões de topografia complexa, onde a variação de altitude e a orientação das encostas afetam diretamente a resposta

espectral das culturas (Moreira *et al.*, 2016; Sovann *et al.*, 2025). Inserida nesse contexto, a região das Matas de Minas, localizada no estado de Minas Gerais e detentora de Indicação Geográfica de Procedência desde 2020, constitui um território de destaque para a cafeicultura nacional. Com uma extensão de 17.633,18 km² distribuída por 64 municípios, caracteriza-se por relevo montanhoso e fortemente ondulado, o que impõe desafios adicionais ao mapeamento automatizado das lavouras (INPI, 2020b).

Neste cenário, o uso de imagens de sensoriamento remoto de alta resolução, como as fornecidas pelo PlanetScope, que possuem resolução espacial que variam entre 3 e 5 metros, cobertura temporal quase diária e quatro bandas espectrais, são elas: azul, verde, vermelho e infravermelho próximo, essas imagens permitem capturar informações detalhadas sobre a vegetação e diferentes tipos de cobertura do solo (Roy *et al.*, 2021). Segundo Picoli *et al.* (2020) tais características viabilizam tanto o monitoramento contínuo das lavouras ao longo de suas fases fenológicas quanto a identificação de talhões individuais, mesmo em propriedades de pequena escala, que são predominantes na região das Matas de Minas.

Diante desse contexto, este trabalho tem como objetivo segmentar lavouras de café na região das Matas de Minas, estado de Minas Gerais, Brasil. Para isso, foram avaliadas diferentes combinações de bandas espectrais e seu impacto no desempenho da rede de segmentação semântica U-Net na identificação dessas lavouras, utilizando imagens de sensoriamento remoto obtidas do PlanetScope.

2 REVISÃO DE LITERATURA

2.1 CAFEICULTURA

O café representa uma das commodities agrícolas mais importantes no cenário econômico global, sendo a segunda mercadoria mais comercializada, atrás apenas do petróleo (PANCSIRA, 2022). Originário do continente africano, o cafeeiro difundiu-se globalmente ao longo dos séculos, consolidando-se como uma cultura essencial tanto do ponto de vista econômico quanto social. Dentre as espécies cultivadas, destacam-se *Coffea arabica* e *Coffea canephora* (Ferreira *et al.*, 2019). A primeira, responsável por cerca de 60% da produção mundial, é cultivada em regiões de maior altitude, com condições climáticas mais amenas, sendo apreciada pela qualidade da bebida. Já a segunda, conhecida como robusta ou conilon, é cultivada em regiões mais quentes e úmidas, representando aproximadamente 40% da produção

global, com destaque para sua rusticidade e resistência a pragas e doenças (Van der Vossen; Bertrand; Charrier, 2015).

No Brasil, o café foi introduzido no início do século XVIII e rapidamente se consolidou como um dos pilares da economia nacional (Topik, 2004). A partir do século XIX, a cafeicultura expandiu-se significativamente, especialmente no Rio de Janeiro, em São Paulo e Minas Gerais, tornando-se a base das exportações brasileiras e desempenhando papel central no processo de modernização do país (Topik, 2019). O ciclo do café contribuiu diretamente para a construção da infraestrutura nacional, incluindo ferrovias, portos e centros urbanos, além de ter impulsionado a imigração europeia como força de trabalho nas lavouras (Faria; Barreira, 2025).

Atualmente, o Brasil é o maior produtor e exportador de café do mundo, responsável por aproximadamente um terço da produção e das exportações globais (Utrilla-Catalan *et al.*, 2022). A atividade concentra-se principalmente nos estados de Minas Gerais, São Paulo, Espírito Santo, Bahia e Paraná, regiões que apresentam condições edafoclimáticas distintas e produzem cafés com ampla diversidade de perfis sensoriais, reconhecidos internacionalmente (Martinez *et al.*, 2024). O cultivo é majoritariamente realizado em propriedades familiares, mas abrange também grandes empreendimentos agrícolas, que variam desde sistemas tradicionais até lavouras altamente tecnificadas, com crescente adoção de práticas de agricultura de precisão (Sott *et al.*, 2020). Além de sua importância econômica direta, a cafeicultura brasileira exerce influência significativa em toda a cadeia produtiva, abrangendo desde a produção primária até a torrefação, exportação e comercialização no varejo, configurando-se como uma das atividades agroindustriais mais relevantes para a geração de empregos no país (Correia *et al.*, 2024).

Apesar de sua posição de liderança no mercado mundial, a cafeicultura brasileira enfrenta desafios crescentes e complexos. A volatilidade dos preços internacionais, os custos elevados de produção, a maior incidência de pragas e doenças e, sobretudo, os impactos das mudanças climáticas colocam em risco a estabilidade da produção e a adequação das áreas de cultivo (INTERNATIONAL COFFEE ORGANIZATION, 2021; Valérie *et al.*, 2024; Verburg *et al.*, 2019). Neste contexto, o uso de tecnologias de sensoriamento remoto, técnicas de agricultura de precisão e ferramentas de inteligência artificial tem se mostrado cada vez mais relevante para o monitoramento e gestão das lavouras cafeeiras (Wang *et al.*, 2024). Portanto, a adoção dessas tecnologias é indispensável para aumentar a resiliência do setor e manter sua competitividade no mercado global.

2.2 SENSORIAMENTO REMOTO

Sensoriamento remoto pode ser definido como a técnica que permite a obtenção de informações sobre alvos ou áreas de interesse, por meio da detecção e análise da radiação eletromagnética refletida ou emitida, isto sem a necessidade de contato físico direto com o objeto observado (DeFries, 2013; Martin, 2008). A radiação eletromagnética compreende todas as ondas eletromagnéticas existentes, que são classificadas conforme seus comprimentos de onda. Segundo Queiroz *et al.* (2022), no sensoriamento remoto, utilizam-se principalmente as faixas do ultravioleta (300 a 400 nm), do espectro visível (400 a 700 nm), do infravermelho próximo (700 a 1.300 nm), do infravermelho médio (1.300 a 5.600 nm) e do infravermelho longo (5.600 a 15.000 nm). Além dessas, também se empregam micro-ondas e ondas de rádio, que possuem comprimentos de onda superiores, sendo as ondas de rádio acima de 27.000 nm (Queiroz *et al.*, 2022). Cada faixa do espectro possui características específicas de interação com a matéria, fornecendo diferentes tipos de informações conforme a aplicação e o sensor utilizado.

Essa detecção da radiação eletromagnética pode ocorrer de duas formas principais: por meio de sensores passivos ou ativos (Marble; Peuquet, 2017). Sensores passivos não emitem radiação própria, eles apenas captam a radiação refletida pelos objetos na superfície, sendo a luz solar a principal fonte (Javaid *et al.*, 2021). Equipamentos como radiômetros e espectrômetros são exemplos comuns, operando em diferentes faixas do espectro eletromagnético (Crewell; Wendisch; Löhnert, 2021). Esse tipo de sensoriamento é utilizado para monitorar a vegetação (Anderson *et al.*, 2016), a temperatura da superfície terrestre e dos oceanos (Duan *et al.*, 2020; Embury *et al.*, 2024), além de analisar a composição da atmosfera, incluindo a presença de nuvens e aerossóis (Islam *et al.*, 2017). Apesar de sua importância, os sensores passivos têm desempenho limitado em áreas com cobertura constante de nuvens densas (Tapakis; Charalambides, 2013).

Por outro lado, o sensoriamento remoto ativo envolve o uso de sensores que emitem sua própria fonte de radiação, geralmente na forma de pulsos de micro-ondas ou luz laser, e medem a energia refletida ou retroespalhada pelos objetos de interesse (Zhang; Li, 2023). Isso permite a coleta de dados independentemente da iluminação solar e mesmo em condições atmosféricas adversas, como chuva ou cobertura de nuvens (Winker *et al.*, 2017). Tecnologias como o RADAR (Radio Detection and Ranging), SAR e o LiDAR (Light Detection and Ranging) são exemplos de sistemas ativos, capazes de medir com precisão a distância, a velocidade e a forma

dos alvos a partir do tempo decorrido entre a emissão e o retorno do sinal (Javaid *et al.*, 2021). Esses sensores são utilizados para mapear topografia, medir a estrutura de florestas, detectar mudanças na superfície terrestre, monitorar ventos, gelo e precipitação, além de obter perfis verticais detalhados da atmosfera. A capacidade de operação independente das condições de luz e clima torna o sensoriamento ativo uma ferramenta estratégica para diversas aplicações ambientais e geoespaciais.

Cada tipo de superfície pode interagir com a radiação eletromagnética de maneira única, refletindo, absorvendo ou transmitindo diferentes quantidades de energia em distintos comprimentos de onda. Essa resposta espectral específica, chamada de assinatura espectral, permite identificar e diferenciar materiais com base em seus padrões de reflectância ao longo do espectro (Queiroz *et al.*, 2022). Como a vegetação, que tende a refletir mais radiação na região do infravermelho e a absorver radiação no espectro do visível. Essa absorção na região do visível, principalmente nos comprimentos de onda do azul e do vermelho ocorre pelo fato dos pigmentos foliares absorverem para a realização da fotossíntese. No entanto, diferentes tipos de plantas podem apresentar variações significativas nesse comportamento espectral, refletindo e absorvendo a radiação de forma distinta conforme suas características fisiológicas, estágio de desenvolvimento, estado nutricional e nível de estresse hídrico ou ambiental (Iqbal *et al.*, 2021; Manjunath; Ray; Panigrahy, 2011).

2.3 MACHINE LEARNING

A inteligência artificial é definida como a capacidade de máquinas, como computadores e robôs, de realizar tarefas que normalmente exigiriam inteligência humana (Chang *et al.*, 2018). Seus fundamentos surgiram na década de 1940, quando McCulloch; Pitts, (1943) propuseram um modelo computacional baseado no comportamento de neurônios biológicos. Embora funcional para operações simples, o modelo apresentava limitações, como saídas binárias e parâmetros fixos (Muthukrishnan *et al.*, 2020). Em 1950, Alan Turing contribuiu para o avanço da área com o artigo “Computing Machinery and Intelligence”, onde propôs o famoso Teste de Turing como forma de avaliar a capacidade de inteligência dessas máquinas.

O conceito de *Machine Learning* emergiu diretamente dessas ideias iniciais. Em 1947, Turing já mencionava a importância de criar máquinas capazes de aprender com a experiência (Edwards; Kaplan; Jie, 2021). O *Machine Learning* é um ramo da Inteligência Artificial voltado à construção de modelos computacionais capazes de identificar padrões e realizar previsões

(Shinde; Shah, 2018). Esses modelos aprendem a partir de dados e ajustam seu desempenho com base em erros anteriores, de forma autônoma (Sarker, 2021), buscando replicar, por meio de algoritmos, a habilidade humana de reconhecer e interpretar padrões complexos (Greener *et al.*, 2022).

As técnicas de *Machine Learning* são geralmente organizadas em quatro categorias principais: supervisionado, não supervisionado, semi-supervisionado e por reforço (Butler *et al.*, 2018). No aprendizado supervisionado, modelos são treinados com dados rotulados para realizar previsões (Zhou *et al.*, 2022). No não supervisionado, o algoritmo trabalha com dados não rotulados e busca identificar padrões ou agrupamentos ocultos (HAN; PEI; Tong, 2022). A abordagem semi-supervisionada combina ambos os tipos de dados, permitindo melhor desempenho com menos exemplos rotulados (Huynh; Nibali; He, 2022). Já o aprendizado por reforço envolve agentes que tomam decisões em ambientes dinâmicos, aprendendo com recompensas ou penalidades, e é amplamente utilizado em tarefas como controle de robôs, jogos e veículos autônomos (Matsuo *et al.*, 2022; Zhou *et al.*, 2022).

Já o *Deep Learning*, é um subcampo do *Machine Learning* e da Inteligência Artificial, utiliza *Deep Artificial Neural Networks* compostas por múltiplas camadas ocultas para aprender automaticamente padrões complexos a partir de grandes volumes de dados, permitindo previsões precisas e decisões eficazes (Zuo *et al.*, 2022). Este conceito foi introduzido por Hinton; Salakhutdinov, (2006), o que impulsionou uma nova área de pesquisa com um dos impactos mais importantes na ciência da computação (Shrestha; Mahmood, 2019). As *Deep Neural Networks* operam em diferentes níveis de abstração e incluem técnicas como Redes Neurais Recorrentes e Redes Neurais Convolucionais, projetadas para explorar relações em dados e gerar saídas confiáveis mesmo em contextos complexos (Lecun; Bengio; Hinton, 2015; Yasaka *et al.*, 2018).

2.4 SEGMENTAÇÃO SEMÂNTICA

Antes do avanço nas técnicas em *Deep Learning*, algoritmos como o *Random Forest* e *Conditional Random Fields* eram amplamente utilizadas para a construção de classificadores em tarefas de aprendizado semântico (Mo *et al.*, 2022). Essas abordagens, embora úteis, apresentavam limitações em capturar padrões visuais complexos. Com o surgimento e consolidação do *Deep Learning*, especialmente por meio das Redes Neurais Convolucionais,

houve uma transformação significativa nesse cenário, elevando o desempenho dos modelos e tornando essas redes a principal escolha em desafios relacionados à segmentação de imagens.

Entre esses desafios, destaca-se a segmentação semântica, que é uma tarefa central na visão computacional, que tem se beneficiado diretamente dos avanços em *Deep Learning*. Ela consiste em rotular cada pixel de uma imagem em uma respectiva classe, permitindo a identificação precisa de diferentes elementos, como ruas, vegetação ou edificações (Long; Shelhamer; Darrell, 2015). Arquiteturas como Fully Convolutional Networks e U-Net tornaram possível extrair representações espaciais e contextuais mais profundas e precisas, superando métodos convencionais baseados em regras fixas ou atributos manuais (Ronneberger; Fischer; Brox, 2015a). Isso ampliou consideravelmente o potencial de aplicação da segmentação semântica em áreas como agricultura de precisão, veículos autônomos, medicina diagnóstica e sensoriamento remoto, consolidando seu papel como uma das aplicações mais impactantes do *Deep Learning* na análise de imagens.

2.5 ESTUDOS RELACIONADOS

2.5.1 Baseados em *Machine Learning*

Técnicas de *Machine Learning* vêm sendo muito exploradas em aplicações de sensoriamento remoto voltadas ao mapeamento agrícola, em especial na classificação de uso e cobertura da terra. Algoritmos como *Random Forest* e Redes Neurais Convolucionais têm sido utilizados para processar imagens de satélite e extrair informações sobre culturas específicas, incluindo lavouras de café. Esses métodos são aplicados em diferentes abordagens e contextos, com variações quanto aos tipos de dados utilizados e aos objetivos de classificação.

O *Random Forest* tem sido empregado em estudos de classificação de uso da terra e análise de lavouras de café. Entre esses, destacam-se o trabalho de Bourgoin *et al.* (2020b) que utilizaram o algoritmo *Random Forest* para classificar o uso e cobertura do solo a partir de bandas espectrais e índices de vegetação. Com índice Kappa variando de 0,72 (1973) a 0,95 (2018). Já Chemura; Mutanga, (2017) utilizaram *Random Forest* na classificação de imagens de satélite, que exploraram o potencial dessa abordagem para mapear classes etárias da cultura do café utilizando dados do Landsat 8 e Landsat 7. O estudo demonstrou que o *Random Forest*, aliado as imagens do Landsat 8, foi capaz de discriminar com precisão entre plantas de café jovem, adulto e velho, com acurácia geral de até 0,903. A análise incluiu a avaliação da

importância das bandas espectrais, destacando a infravermelho próximo como a mais relevante para a separação entre as classes de idade, especialmente no Landsat 8.

Além do uso do *Random Forest*, abordagens baseadas em *Deep Neural Networks* também têm sido exploradas na classificação de áreas agrícolas. Nesse contexto, Baeta *et al.* (2017) propuseram uma abordagem de segmentação semântica pixel a pixel baseada na combinação de Redes Neurais Convolucionais operando em múltiplas escalas (17×17, 25×25 e 33×33 pixels). A estratégia multiescala alcançou acurácia geral de 0,889. Destacando-se também na detecção de lavouras de café recém-plantadas e na identificação de áreas não cultivadas entre talhões. Apresentou também valores baixos de falsos negativos e mostrou ser eficiente ao lidar com áreas de vegetação densa e cenários montanhosos.

2.5.2 Baseados em Classificador da Máxima Verossimilhança

O Classificador de Máxima Verossimilhança é um método estatístico empregado na classificação de pixels em imagens. Fundamentado na teoria bayesiana, o Classificador de Máxima Verossimilhança assume que os dados de cada classe seguem uma distribuição normal multivariada, e, a partir disso, calcula a probabilidade de cada pixel pertencer a uma determinada classe com base nas estatísticas extraídas de amostras (Otukey; Blaschke, 2010). O pixel é então rotulado com a classe cuja função de verossimilhança apresenta o maior valor. Essa abordagem considera tanto a posição média das classes no espaço espectral quanto a dispersão dos dados, o que o torna mais robusto do que classificadores baseados apenas na distância. Por sua capacidade de incorporar a variabilidade interna das classes, o Classificador de Máxima Verossimilhança é eficaz na discriminação de alvos espectralmente semelhantes, como diferentes tipos de vegetação em um mesmo bioma, sendo, por isso, uma escolha frequente em estudos de uso e cobertura da terra (Novo, 1992).

Estudos demonstram o uso do Classificador de Máxima Verossimilhança na identificação e mapeamento de áreas de cultivo de café. (Schmitt-Harsh, 2013b) Utilizaram diferentes combinações de bandas espectrais para treinar o Classificador de Máxima Verossimilhança para gerar cinco classes de cobertura do solo, que inclui floresta, agrofloresta de café, pastagem, plantações de banana e seringueira e assentamentos rurais. Já Lelong; Thong-Chane, (2003) utilizaram imagens pancromáticas dos satélites Ikonos e Quickbird para realizar a classificação de diversas culturas, incluindo duas categorias de café. A classificação por máxima verossimilhança foi testada para diferentes conjuntos de índices texturais. E

Gaertner *et al.* (2017) utilizaram o Classificador de Máxima Verossimilhança para classificar o café no Havaí com base em oito bandas espectrais.

2.5.3 Utilização de sensores ativos na identificação de lavouras de café

Sensores ativos, como os SARs, operam de forma independente das condições meteorológicas e da iluminação solar, fornecendo imagens a partir da emissão e recepção do próprio sinal. Essa característica permite a aquisição regular de dados ao longo do tempo, o que tem motivado a realização de estudos voltados à aplicação de diferentes configurações de polarização SAR na identificação e distinção de culturas agrícolas. Entre essas culturas, o café tem sido objeto de investigação, considerando suas particularidades estruturais e padrões de plantio que podem influenciar a resposta do sinal.

Entre os estudos voltados a essa abordagem, Silva *et al.* (2009b) investigaram o uso de dados SAR aerotransportados (sensor SAR-R99B, banda L) para a classificação de culturas agrícolas em ambiente tropical, com ênfase na cafeicultura. Foram testadas diferentes configurações de polarização (simples, dupla, tripla e polarimétrica completa), evidenciando um aumento expressivo na precisão conforme a complexidade dos dados. As polarizações simples apresentaram resultados limitados, com acurácia geral variando entre 0,384 e 0,564, enquanto os dados totalmente polarimétricos alcançaram índices superiores a 0,96. A distinção entre classes de café foi favorecida por características estruturais específicas das plantações, como o espaçamento entre linhas e a direção do plantio, que influenciaram a resposta do sinal SAR. O uso combinado de diferentes polarizações, especialmente em conjunto com o classificador ICM, permitiu uma segmentação mais eficaz das áreas de café em comparação com outras culturas, como algodão e pastagens.

3 MATERIAL E MÉTODOS

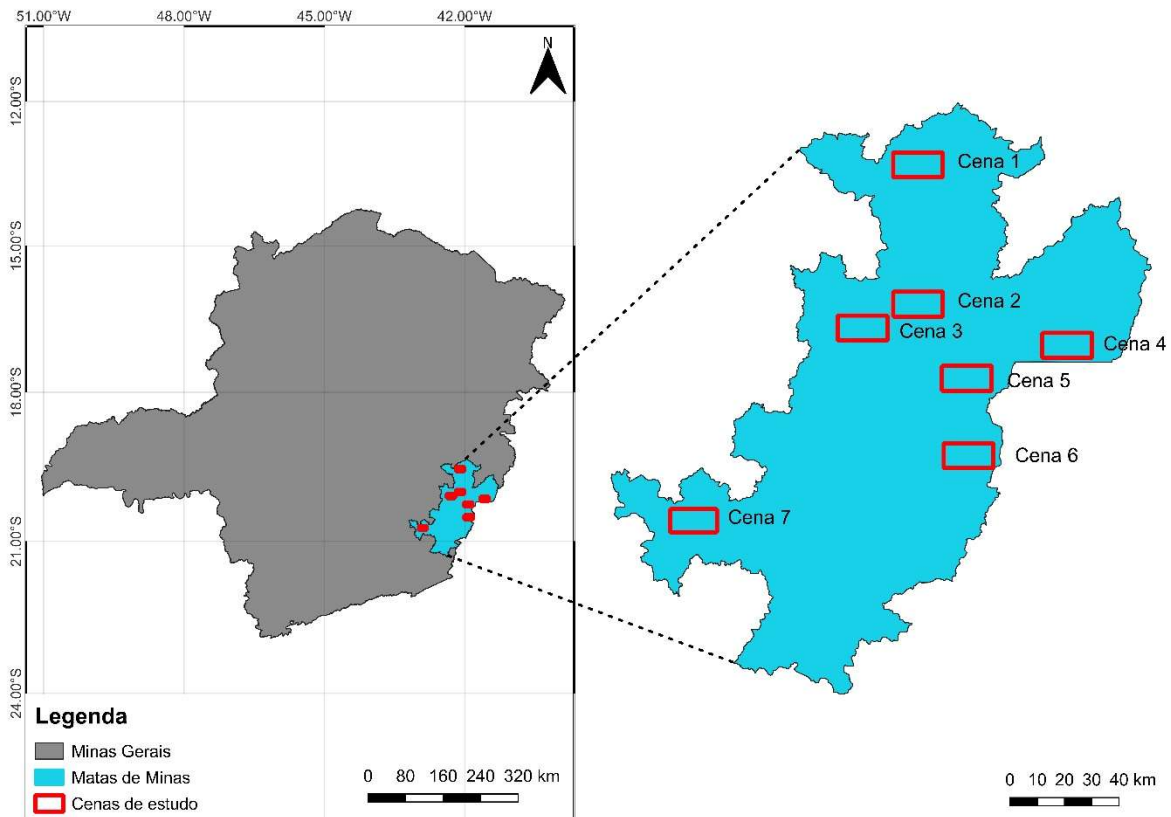
3.1 ÁREA DE ESTUDO E IMAGENS UTILIZADAS

O estudo foi realizado na região denominada Matas de Minas, estado de Minas Gerais, Brasil, conforme indicado na Figura 1. Esta região recebeu a concessão da Indicação Geográfica do tipo “Indicação de Procedência” no ano de 2020. A Indicação Geográfica (IG) é um registro concedido pelo Instituto Nacional da Propriedade Industrial (INPI) que reconhece oficialmente a reputação ou a qualidade de um produto ou serviço vinculada à sua origem geográfica,

podendo ser classificada como Indicação de Procedência (IP) ou Denominação de Origem (DO) (INPI, 2020). Além disso, a região abrange uma extensão de 17.633,18 km², compreendendo 64 municípios e é caracterizada por um relevo montanhoso e fortemente ondulado. Grande parte do território possui altitudes que variam entre 400 e 1.000 metros, porém, existem locais que ultrapassam os 2.000 metros (INPI, 2020a).

Dentro dessa área, foram selecionadas sete cenas, a partir das quais foram obtidas diversas imagens de sensoriamento remoto do PlanetScope, que possuem resolução espacial de quatro metros e quatro bandas espectrais: azul (455-515 nm), verde (500-590 nm), vermelho (590-670 nm) e infravermelho próximo (780-860 nm) (Roy *et al.*, 2021). Ao todo, foram selecionadas 54 imagens distribuídas entre as sete cenas, correspondendo a registros mensais referentes aos anos de 2021, 2022, 2023 e 2024, escolhendo aquelas com baixa cobertura de nuvens, de modo a não comprometer a identificação visual das classes de cobertura do solo. Essas 54 imagens estão distribuídas entre as sete cenas selecionadas na região, conforme apresentado na Figura 1.

Figura 1. Localização da área de estudo. A região em azul corresponde às Matas de Minas, enquanto as áreas em vermelho destacam as sete cenas selecionadas (Cenas 1 a 7). Dessas cenas foram obtidas 54 imagens mensais do sensor PlanetScope, adquiridas entre 2021 e 2024.



Fonte: Do autor

3.2 DESENVOLVIMENTO DO MODELO

O treinamento do modelo foi conduzido em ambiente de alto desempenho computacional, utilizando o cluster da Universidade Federal de Viçosa (UFV). O nó principal (head node) do cluster é equipado com um processador Intel Xeon Gold 6212U, com 24 núcleos e 48 threads, operando a 2,40 GHz, além de 512 GB de memória RAM DDR4 a 3200 MT/s. Esse ambiente de alto desempenho foi utilizado para coordenar as tarefas computacionais, oferecendo os recursos necessários para o processamento eficiente dos dados e execução dos modelos.

Para o desenvolvimento do modelo, foram utilizadas imagens mensais das cenas 2 e 7 no processo de treinamento e validação, enquanto para as demais cinco cenas, reservadas para o teste, foi utilizada apenas uma imagem por cena. Para a fase de treinamento foi utilizado o seguinte fluxo de trabalho: rotulação das classes nas imagens, pré-processamento dos dados e modelagem com as diferentes combinações de bandas espectrais. Após a definição do melhor modelo, obtido a partir da combinação de bandas espectrais que apresentou o melhor desempenho no conjunto de validação, foi realizada a etapa de teste utilizando exclusivamente imagens das cenas que não participaram do treinamento ou da validação. Para a implementação da arquitetura U-Net, foi utilizada a linguagem *Python* na versão 3,13.

3.2.1 Rotulagem

Com base nas imagens do PlanetScope, referentes às cenas 2 e 7, foi realizada a rotulagem manual das classes. Foram definidas quatro classes de interesse: “Coffee Crops”, “Forest”, “City” e “Other”. A classe Coffee Crops corresponde a áreas cultivadas com café. A classe Forest inclui regiões com cobertura florestal densa, tanto nativa quanto de reflorestamento. A classe City abrange áreas urbanas, casas, construções rurais e demais estruturas artificiais. Por fim, a classe Other engloba todas as demais superfícies, como pastagens, corpos d’água, solo exposto e culturas agrícolas distintas do café, que não apresentaram características visuais compatíveis com as classes principais.

A escolha por quatro classes de interesse, em vez da simplificação em apenas “café” e “não café”, teve como objetivo lidar com a heterogeneidade da paisagem e minimizar a confusão espectral, sobretudo em áreas com características semelhantes às lavouras de café. A definição dessas classes considerou a capacidade de distinguir visualmente cada uma delas nas imagens do PlanetScope, o que permitiu maior consistência na rotulagem. Para compor o

conjunto de dados de treinamento e validação, foram selecionadas as cenas 2 e 7 (Figura 1), como área de estudo.

Para as cenas 2 e 7, foram obtidas imagens mensais referentes aos anos de 2021 a 2024, totalizando 49 imagens, sobre as quais se aplicou uma abordagem de rotulagem densa, conforme metodologia abordada por Tong; Xia; Zhu, (2023). Essa técnica consiste na criação de múltiplos polígonos que cobrem sub-regiões menores com alto nível de detalhamento, permitindo maior precisão na delimitação das classes de interesse. Os polígonos foram elaborados no QGIS (versão 3.22.1), no formato shapefile e no sistema de coordenadas WGS 84, o mesmo utilizado nas imagens.

Com base nas rotulagens manuais realizadas, cada imagem foi analisada visualmente e as diferentes classes de cobertura do solo foram delimitadas por meio da criação de múltiplos polígonos. Essa abordagem de rotulagem densa permitiu classificar cada pixel de acordo com a sua classe correspondente “Coffee Crops”, “Forest”, “City” ou “Other”, garantindo maior representação entre as classes. Como resultado, o conjunto de dados de treinamento e validação apresentou um total de 1.264.844.800 pixels classificados, conforme demonstrado na Tabela 1.

Tabela 1. Distribuição de pixels rotulados por classe no conjunto de dados para treinamento e validação

Coffee Crops	Forest	City	Other	Total
89.787.927	298.548.871	51.860.820	824.647.182	1.264.844.800

Para as cinco cenas independentes (cenas 1, 3, 4, 5 e 6), destinadas à composição do conjunto de teste, foi obtida uma imagem para cada cena, totalizando cinco imagens. Nessas cenas, aplicou-se a mesma metodologia de rotulagem densa utilizada no treinamento e validação. Cada imagem foi analisada visualmente no QGIS (versão 3.22.1), e as classes de cobertura do solo foram delimitadas por meio da criação de múltiplos polígonos em formato shapefile, seguindo o sistema de coordenadas WGS 84. Essa abordagem permitiu a rotulagem das classes "Coffee Crops", "Forest", "City" e "Other", assegurando consistência metodológica entre os conjuntos de dados e maior precisão na avaliação do modelo em áreas geograficamente distintas.

3.2.2 Pré-processamento dos dados

Inicialmente, o conjunto de dados destinado a treinamento e validação foi particionado de forma a garantir a separação total entre os subconjuntos utilizados para treinamento e validação. Primeiramente, 44 imagens foram reservadas para o conjunto de treinamento e cinco imagens para o conjunto de validação, assegurando que não houvesse qualquer sobreposição entre eles.

Embora as imagens para treinamento e validação tenham sido extraídas das mesmas cenas (cenas 2 e 7), foram utilizadas imagens obtidas em períodos distintos, com datas de aquisição variando entre diferentes meses e anos (2021 a 2024). Para a validação, foram selecionadas cinco imagens específicas de meses e anos variados, de modo a minimizar possíveis vieses. Essa estratégia visou reduzir tanto a sobreposição espectral entre os conjuntos quanto a exposição do modelo a padrões espectrais semelhantes durante o treinamento e a validação.

Os resultados obtidos na validação foram utilizados como referência para a seleção do melhor modelo. As imagens de validação vieram das mesmas cenas de treinamento, no entanto, como são de períodos diferentes, existem modificações das áreas cultivadas. Assim, foi possível avaliar se um modelo desenvolvido para uma determinada cena atualiza com as lavouras de café em produção.

As imagens no formato TIFF e os arquivos de anotação no formato shapefile foram carregados e organizados em lotes para otimizar o uso de memória. Foram avaliadas diferentes combinações de bandas espectrais como entrada no modelo, porém confinados a três bandas espectrais combinadas (B, G, R), (B, G, NIR), (B, R, NIR) e (G, R, NIR). As siglas B, G, R e NIR se referem às bandas espectrais azul, verde, vermelho e infravermelho próximo, respectivamente. Para cada imagem, foram geradas máscaras a partir da rotulagem, onde as classes foram codificadas como 1 (Coffee Crops), 2 (Forest), 3 (City) e 0 (Other).

Após a separação entre os conjuntos de treinamento e validação, as imagens foram então particionadas em patches de 256x256 pixels, ou seja, em pequenas subimagens extraídas da imagem principal. Esse particionamento resultou em um total de 19.300 patches de 256x256 pixels extraídos das 49 imagens das cenas 2 e 7. Esse procedimento reduz a complexidade computacional e melhora a eficiência no treinamento do modelo. Após essa etapa, os valores dos pixels foram normalizados para o intervalo [0, 1], uniformizando as entradas para a rede neural. Além disso, foi utilizado um gerador de dados simples, implementado para fornecer os

patches em lotes. Sua função principal é otimizar o uso de memória e garantir o fornecimento contínuo de amostras durante o processo de treinamento e validação.

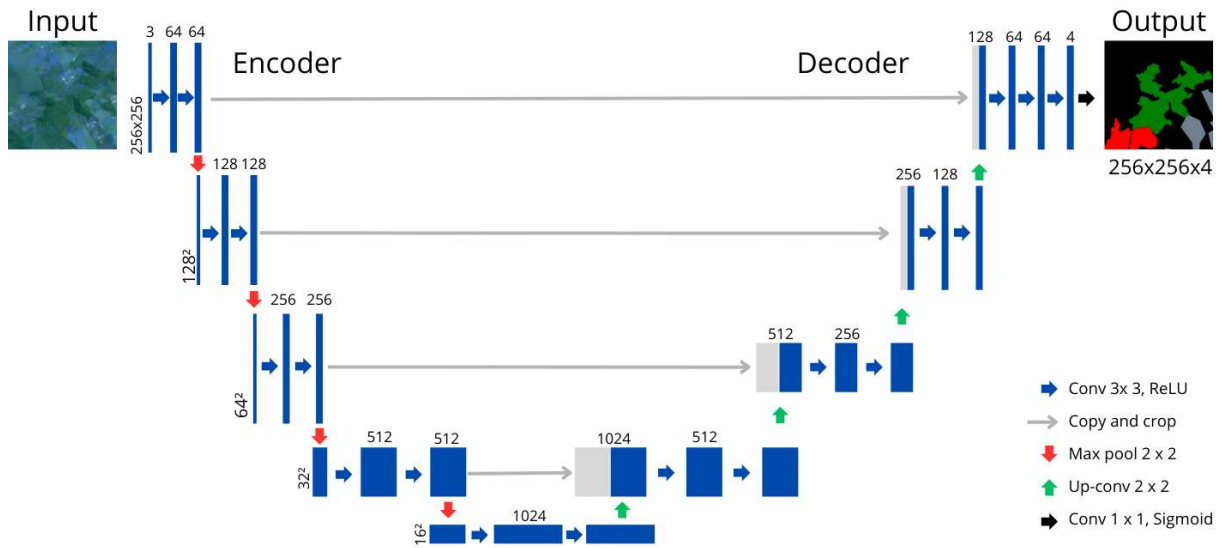
3.2.3 Modelagem

Neste estudo, utilizou-se a rede de segmentação semântica U-Net (Ronneberger; Fischer; Brox, 2015a) para construir o modelo de identificação de lavouras de café. O modelo foi treinado por 10 épocas para cada combinação de bandas espectrais, utilizando o otimizador *Adam* com taxa de aprendizado inicial de 0,0001, função de perda *categorical crossentropy* e *early stopping* com paciência de cinco épocas para evitar *overfitting*. Durante o treinamento, foi implementado um *callback* para redução da taxa de aprendizado caso a perda estagnasse por cinco épocas consecutivas. O modelo com melhor desempenho foi salvo e utilizado para avaliação no conjunto de validação e testes nas cenas independentes.

A arquitetura utilizada neste trabalho segue os princípios estruturais da U-Net, composta por um caminho codificador (*encoder*) e um caminho decodificador (*decoder*). O *encoder* é responsável pela extração de características, com quatro blocos sucessivos contendo duas camadas de convolução seguidas de operações de *max pooling*. O *decoder*, por sua vez, tem como função recuperar a resolução da imagem, utilizando operações de *upsampling* seguidas por duas camadas de convolução. As conexões do tipo *skip* interligam os blocos correspondentes do *encoder* ao *decoder*, permitindo a preservação de informações espaciais ao longo da rede.

A Figura 2 apresenta a arquitetura U-Net utilizada neste trabalho, construída e adaptada da versão original proposta por Ronneberger; Fischer; Brox, (2015a). Diferentemente da versão original, cuja estrutura utiliza imagens de entrada com dimensão de 572×572 pixels e gera saídas com 388×388 pixels. A rede aqui implementada foi ajustada para receber patches de 256×256 pixels, compatível com os patches das imagens utilizadas, com três bandas espectrais e saída configurada para quatro classes. Essa adaptação assegura a adequação da rede à tarefa proposta, mantendo os princípios fundamentais de organização estrutural, profundidade e fluxo de informações característicos da U-Net.

Figura 2. Arquitetura U-Net adaptada e utilizada neste trabalho. O *Input* corresponde a patches de 256×256 pixels com três bandas espectrais, e o *Output* apresenta a máscara de segmentação com quatro classes: Coffee Crops (vermelho), Forest (verde), City (cinza) e Other (preto).



Fonte: Adaptado de Ronneberger; Fischer; Brox, (2015b)

3.2.4 Métricas de Avaliação

A avaliação do modelo foi realizada através da análise do desempenho da arquitetura U-Net na segmentação no conjunto de validação, utilizando as diferentes combinações de bandas espectrais (B, G, R), (B, G, NIR), (B, R, NIR) e (G, R, NIR). As métricas adotadas para avaliação incluíram: Acurácia, que representa a proporção total de pixels classificados corretamente; F1-score, uma medida balanceada que combina Precisão e Recall para indicar a relação entre falsos positivos e falsos negativos; e Intersecção sobre União (IoU), que mede a área de intersecção entre a predição e a classe verdadeira dividida pela área de união, sendo uma métrica amplamente usada para quantificar a qualidade da segmentação. As fórmulas para calcular a Acurácia, F1-score, Precisão, Recall e IoU são mostradas nas equações:

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$\text{F1 - score} = 2 \times \frac{\text{Precisão} \times \text{Recall}}{\text{Precisão} + \text{Recall}} \quad (2)$$

$$\text{Precisão} = \frac{TP}{TP + FP} \quad (3)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

$$\text{IoU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \quad (5)$$

Nas equações, TP representa o número de pixels corretamente identificados como pertencentes à classe positiva; TN representa o número de pixels corretamente identificados como não pertencentes à classe positiva; FP representa o número de pixels incorretamente identificados como pertencentes à classe positiva; e FN representa o número de pixels incorretamente identificados como não pertencentes à classe positiva.

As métricas de IoU foram calculadas para cada classe (Coffee Crops, Forest, City e Other), e também foi considerada a IoU média entre todas as classes. O desempenho do modelo foi priorizado com base na IoU da classe Coffee Crops no conjunto de validação, por ser a classe de interesse. Além disso, matrizes de confusão foram geradas para cada combinação de bandas, permitindo assim uma análise detalhada da distribuição de erros e do desempenho do modelo nos diferentes cenários.

Após identificar a melhor combinação de bandas espectrais com base no conjunto de validação, o modelo final foi aplicado nas cinco cenas independentes (cenas 1, 3, 4, 5 e 6), estas que não participaram do treinamento nem da validação. Para cada uma dessas imagens, foram gerados mapas de segmentação completos, no qual foram comparadas com as rotulagens manuais realizadas. A comparação entre as predições do modelo e as máscaras de referência foi realizada por meio das métricas estatísticas apresentadas nas Equações (1) a (5), permitindo quantificar o desempenho do modelo.

4 RESULTADOS

A rede de segmentação semântica U-Net foi treinada para a identificação das classes Coffee Crops, Forest, City e Other, utilizando imagens multiespectrais do PlanetScope. A rede foi treinada com as seguintes combinações de bandas espectrais (B, G, R), (B, G, NIR), (B, R, NIR) e (G, R, NIR), buscando analisar a influência da combinação espectral no desempenho da segmentação. O desempenho da rede foi avaliado por meio das métricas IoU, Acurácia, F1-score, Precisão e Recall. Embora todas as classes tenham sido consideradas na avaliação, o foco principal dos experimentos foi analisar como a escolha das bandas espectrais impacta

especificamente o desempenho na segmentação da classe Coffee Crops, por ser o principal alvo do estudo.

4.1 MODELO DESENVOLVIDO

Foi avaliado o desempenho geral da rede U-Net na segmentação de todas as classes do estudo, com base no conjunto de validação, utilizando diferentes combinações de bandas espectrais. Na Tabela 2 são apresentados os resultados das métricas estatísticas quantitativas no desempenho geral obtidas para cada combinação de bandas espectrais testadas.

Tabela 2. Resultados gerais da segmentação da combinação de diferentes bandas com base no conjunto de validação

Combinação de bandas	Acurácia	Precisão	Recall	F1-score	IoU
(B, G, R)	0,9659	0,9683	0,9659	0,9663	0,9354
(B, G, NIR)	0,9496	0,9516	0,9496	0,9499	0,9057
(B, R, NIR)	0,8935	0,9218	0,8935	0,8944	0,8180
(G, R, NIR)	0,9622	0,9645	0,9622	0,9626	0,9287

Os números em negrito indicam que a combinação de bandas espectrais possui o melhor desempenho geral.

Os resultados demonstram que as combinações de bandas (B, G, R) e (G, R, NIR) obtiveram desempenho superior, apresentando os maiores valores para todas as métricas avaliadas (Acurácia, Precisão, Recall, F1-score e IoU). A combinação (B, G, R) alcançou o melhor desempenho geral, com acurácia de 0,9659 e IoU de 0,9354, seguida pela combinação (G, R, NIR) com acurácia de 0,9622 e IoU de 0,9287. Em contraste, a combinação (B, R, NIR) apresentou o menor desempenho entre todas as configurações testadas, com acurácia de 0,8935 e IoU de 0,8180.

Ao analisar o desempenho por classe na segmentação, os valores de F1-score são apresentados na Tabela 3 e demonstram que as combinações (B, G, R) e (G, R, NIR) mantêm superioridade na maioria das classes. A combinação (B, G, R) obteve os melhores resultados nas classes Other (0,9381), Forest (0,9763) e City (0,9908), enquanto (G, R, NIR) teve a maior F1-score em Coffee Crops (0,9616), apresentando desempenho consistente nas demais classes. Já as combinações (B, G, NIR) e (B, R, NIR) apresentaram desempenho inferior, principalmente na classe Coffee Crops, com F1-score de 0,9422 e 0,8114, respectivamente.

Tabela 3. Avaliação da combinação de bandas por classe com base na F1-score e no conjunto de validação

Combinação de bandas	Other	Coffee Crops	Forest	City
(B, G, R)	0,9381	0,9599	0,9763	0,9908
(B, G, NIR)	0,9131	0,9422	0,9585	0,9858
(B, R, NIR)	0,8246	0,8114	0,9522	0,9896
(G, R, NIR)	0,9314	0,9616	0,9692	0,9882

Os números em negrito indicam que a combinação de bandas espectrais possui o melhor desempenho por classe na métrica F1-score.

Complementando essa análise, a Tabela 4 apresenta os valores de IoU por classe, confirmando o padrão observado nas métricas de F1-score.

Tabela 4. Avaliação da combinação de bandas por classe com base na IoU e no conjunto de validação

Combinação de bandas	Other	Coffee Crops	Forest	City
(B, G, R)	0,8834	0,9228	0,9763	0,9817
(B, G, NIR)	0,8401	0,8907	0,9203	0,9720
(B, R, NIR)	0,7016	0,6826	0,9087	0,9793
(G, R, NIR)	0,8716	0,9260	0,9403	0,9766

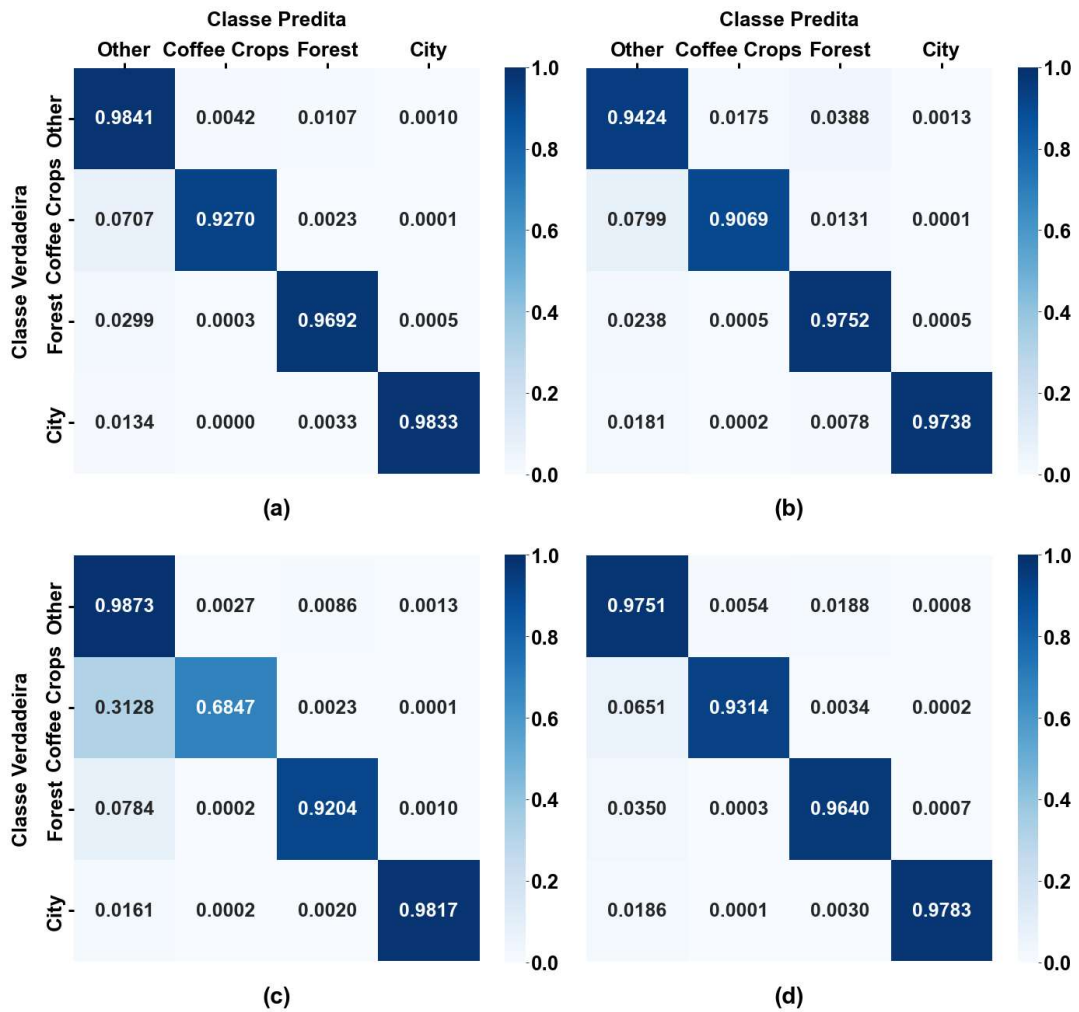
Os números em negrito indicam que a combinação de bandas espectrais possui o melhor desempenho por classe na métrica IoU.

Novamente, as combinações (B, G, R) e (G, R, NIR) se destacam, com desempenho superior na maioria das classes. A combinação (B, G, R) apresentou os maiores valores nas classes Forest (0,9763), City (0,9817) e Other (0,8834), além de resultado alto na classe Coffee Crops (0,9228). A combinação (G, R, NIR) superou todas as demais em Coffee Crops (0,9260) e teve desempenho consistente nas demais classes. Por outro lado, as combinações (B, G, NIR) e (B, R, NIR) apresentaram valores mais baixos, especialmente na classe Coffee Crops, com IoU de 0,8907 e 0,6826, respectivamente. Esses resultados reforçam a eficiência das combinações (B, G, R) e (G, R, NIR) na segmentação por classe, mantendo desempenho robusto em diferentes categorias.

4.2 ANÁLISE DOS PADRÕES DE PREDIÇÃO ENTRE AS CLASSES

Para compreender os erros e o desempenho na métrica IoU, especialmente na classe Coffee Crops, torna-se fundamental a análise das matrizes de confusão, que permitem visualizar de forma precisa como os pixels são classificados de maneira errada. Essa abordagem possibilita a identificação de padrões recorrentes de confusão entre classes específicas, contribuindo para um diagnóstico das limitações da segmentação. As matrizes de confusão foram geradas para cada combinação de bandas espectrais a partir do conjunto de dados de validação, permitindo avaliar o desempenho da segmentação e localizar os principais erros de predição. Na Figura 3 são apresentadas as matrizes de confusão correspondentes a cada combinação de banda como entrada para a arquitetura U-Net.

Figura 3. Matrizes de confusão geradas para cada combinação de bandas espectrais utilizadas como entrada na arquitetura U-Net, avaliadas no conjunto de validação: (a) combinação de bandas (B, G, R); (b) combinação de bandas (B, G, NIR); (c) combinação de bandas (B, R, NIR); e (d) combinação de bandas (G, R, NIR).



Fonte: Do autor

A combinação (B, G, R) apresentou alto desempenho na segmentação das quatro classes, superando 0,9000 de acertos para todas as classes. Destacam-se as classes Other e City,

com taxas de acerto de 0,9841 e 0,9833, respectivamente. A classe Forest alcançou 0,9692, enquanto Coffee Crops apresentou 0,9270. O principal erro ocorreu entre Coffee Crops e Other (0,0707), indicando possível sobreposição de características espectrais que dificultam a distinção pela combinação espectral. As demais confusões foram inferiores a 0,015, o que evidencia a alta precisão da combinação (B, G, R) na maioria dos casos.

Já a combinação (B, G, NIR) também obteve bom desempenho, porém apresentou resultados inferiores a combinação (B, G, R). Forest (0,9752) e City (0,9738) mantiveram altas taxas, similares às da combinação (B, G, R), no entanto, Coffee Crops caiu para 0,9069, contra 0,9270 da (B, G, R), e Other teve 0,9424, bem abaixo dos 0,9841 da (B, G, R). A confusão entre Coffee Crops e Other aumentou para 0,0799, indicando que a banda NIR não ajudou a diferenciar essas classes. No geral, a (B, G, R) foi mais precisa, especialmente para identificar Coffee Crops.

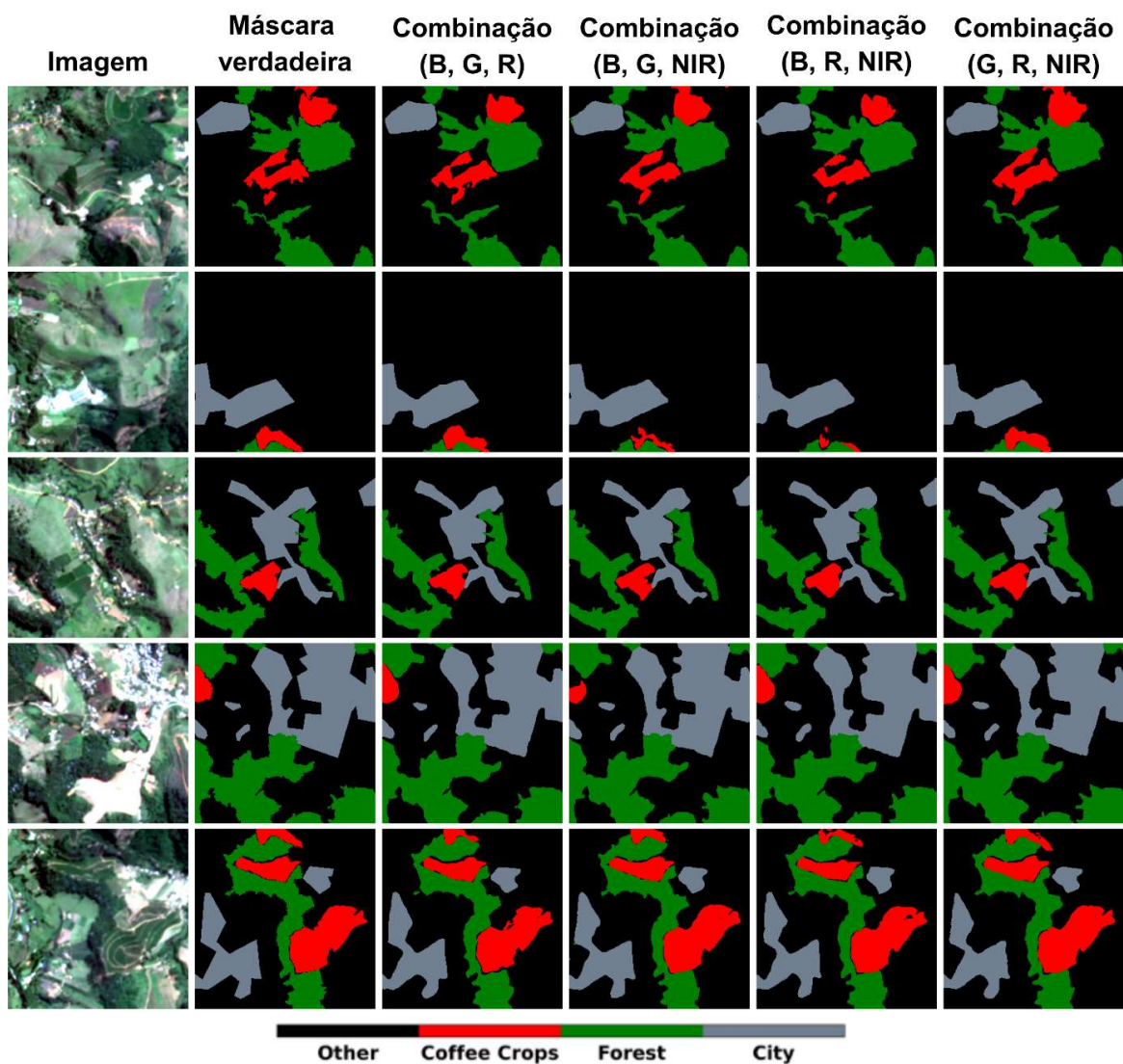
A combinação (B, R, NIR) apresentou desempenho consistente para a maioria das classes, mas inferior às combinações anteriores, especialmente na identificação da classe Coffee Crops. As classes Other e City mantiveram altas taxas de acerto, com 0,9873 e 0,9817, respectivamente, enquanto Forest atingiu 0,9204, também com bom desempenho. No entanto, Coffee Crops teve queda significativa, com apenas 0,6847 de acerto. A principal confusão ocorreu entre Coffee Crops e Other (0,3128), reforçando a dificuldade do modelo em diferenciar essas classes com essa combinação espectral. As demais confusões permaneceram abaixo de 0,01, evidenciando que o problema de classificação foi pontual, concentrado principalmente na distinção entre Coffee Crops e Other.

A combinação (G, R, NIR) apresentou taxas de acerto superiores a 0,93 para todas as categorias. As classes City e Other se destacaram, com 0,9783 e 0,9751, respectivamente. Já a classe Forest apresentou 0,9640 de acerto, enquanto Coffee Crops registrou 0,9314, sendo este o melhor resultado entre todas as combinações espectrais testadas para essa classe. A principal confusão ocorreu entre Coffee Crops e Other, com 0,0651 dos pixels de Coffee Crops classificados como Other, o que evidencia sobreposição espectral entre essas categorias, mesmo com a substituição da banda B pela NIR. As demais confusões foram inferiores a 0,01, indicando que a combinação (G, R, NIR) permite distinguir adequadamente as demais classes.

4.3 ANÁLISE COMPARATIVA DAS MÁSCARAS DE SEGMENTAÇÃO

Por fim, além da análise quantitativa, foi realizada uma avaliação qualitativa das máscaras de segmentação produzidas pelo modelo. A Figura 4 apresenta a imagem que foi segmentada, seguida pela máscara verdadeira, e pelas previsões da U-Net para as quatro combinações espectrais testadas: (B, G, R), (B, G, NIR), (B, R, NIR) e (G, R, NIR), permitindo uma avaliação qualitativa dos resultados obtidos no conjunto de validação.

Figura 4. Resultados da segmentação de imagens do conjunto de validação utilizando as combinações espectrais (B, G, R), (B, G, NIR), (B, R, NIR) e (G, R, NIR), com base na arquitetura U-Net. As cores indicam as classes: Coffee Crops (vermelho), Forest (verde), City (cinza) e Other (preto).



Fonte: Do Autor

Observa-se que as quatro combinações de bandas utilizadas como entrada para a rede U-Net não foram igualmente eficazes na identificação das quatro classes do estudo. Em algumas previsões, as classes foram mal definidas, indicando limitações no desempenho dos modelos

com determinadas combinações espectrais. No entanto, para a classe Coffee Crops, a combinação (G, R, NIR) se destacou, apresentando áreas segmentadas mais coerentes com a máscara verdadeira, com fronteiras bem definidas e pouca confusão com outras classes. Por outro lado, as combinações (B, R, NIR) e (B, G, NIR) apresentaram os piores desempenhos visuais, com falhas evidentes na identificação da classe Coffee Crops, frequentemente confundida com a classe Other.

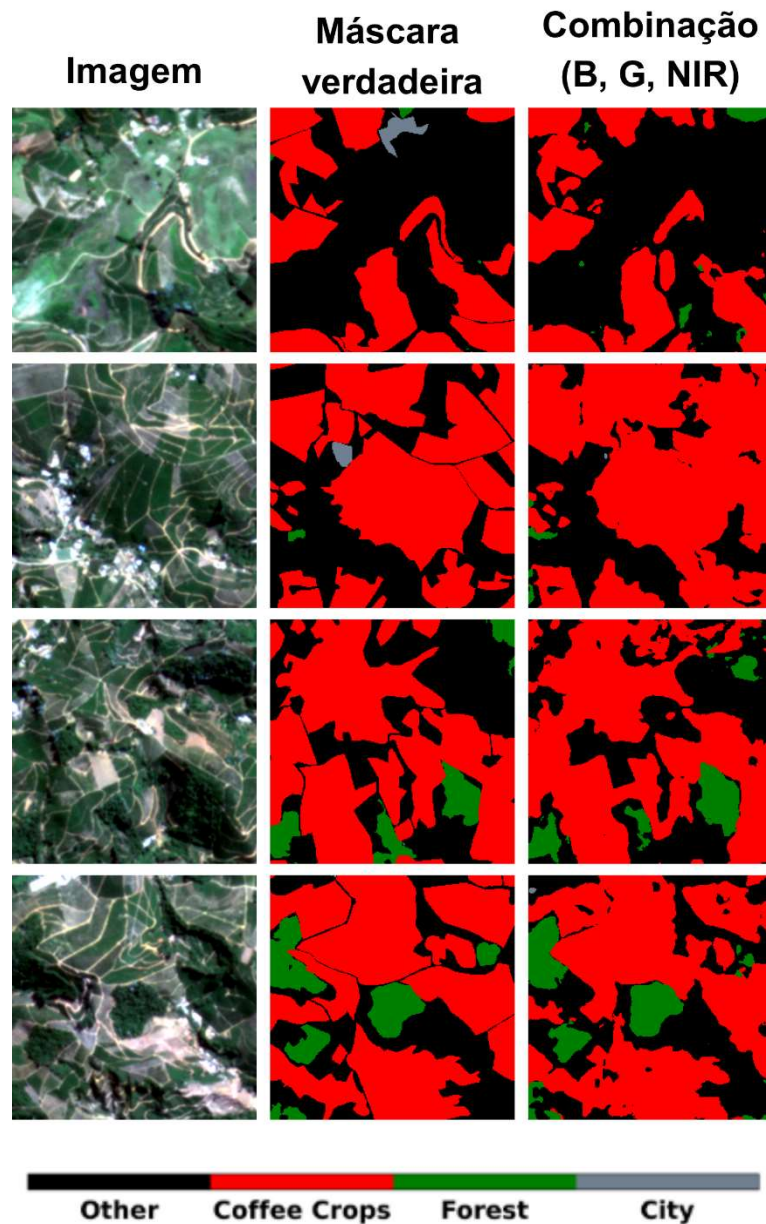
Ao observar as máscaras geradas para cada combinação, percebe-se que os erros mais recorrentes ocorrem nas regiões de borda entre as classes, em especial nas transições entre Coffee Crops e Other. Nas combinações de melhor desempenho, como (B, G, R) e (G, R, NIR), essas transições são, em geral, melhores delimitadas, com falhas pontuais nas margens das áreas das classes Coffee Crops e Other. Já nas combinações de menor desempenho (B, R, NIR) e (B, G, NIR), essas imprecisões nas fronteiras são mais acentuadas, resultando em erros de predição da classe Coffee Crops e resultando em perda de continuidade espacial.

A avaliação visual confirma as tendências observadas nas métricas: as combinações (B, G, R) e (G, R, NIR) produziram segmentações mais consistentes e coerentes, com melhores delimitações de fronteira entre as classes. Já as combinações (B, R, NIR) e (B, G, NIR) apresentaram falhas visíveis, sobretudo na identificação de Coffee Crops, frequentemente confundida com a classe Other.

4.4 TESTE DO MODELO EM NOVAS CENAS

Para avaliar a capacidade de generalização do modelo desenvolvido, foi selecionada a combinação de bandas (G, R, NIR), que demonstrou superior desempenho na segmentação da classe de interesse Coffee Crops com base na métrica IoU no conjunto de validação. Este modelo foi então aplicado em cinco cenas completamente independentes (cenas 1, 3, 4, 5 e 6), que não foram utilizadas em nenhuma etapa do processo de treinamento ou validação. A Figura 5 a seguir apresenta a aplicação do modelo treinado com a combinação (G, R, NIR) nas cenas de teste.

Figura 5. Exemplos de predições do modelo treinado aplicado nas cenas de teste (1, 3, 4, 5 e 6). As cores representam as classes preditas pelo modelo: Other (amarelo), Coffee Crops (vermelho), Forest (verde) e City (cinza).



Fonte: Do Autor

Os resultados gerais obtidos nas cenas de teste (cenas 1, 3, 4, 5 e 6) demonstraram capacidade de generalização do modelo treinado. O modelo apresentou uma acurácia geral de 0,8933, precisão de 0,8962 e o recall de 0,8933, que resultaram em F1-score de 0,8929, indicando equilíbrio entre a capacidade de identificar corretamente as classes e minimizar falsos positivos e negativos. Além disso, o valor de IoU foi de 0,8082, confirmando a qualidade da segmentação por meio de uma sobreposição entre as predições do modelo e as classes verdadeiras.

A análise detalhada por classe, apresentada na Tabela 5, revela o comportamento específico do modelo para cada classe de cobertura do solo nas cenas de teste. Os resultados demonstram variações no desempenho entre as diferentes classes, refletindo suas características espectrais e complexidade de identificação.

Tabela 5. Desempenho por classe do modelo com combinação de bandas (G, R, NIR) nas cinco cenas de teste, com valores de F1-score e IoU

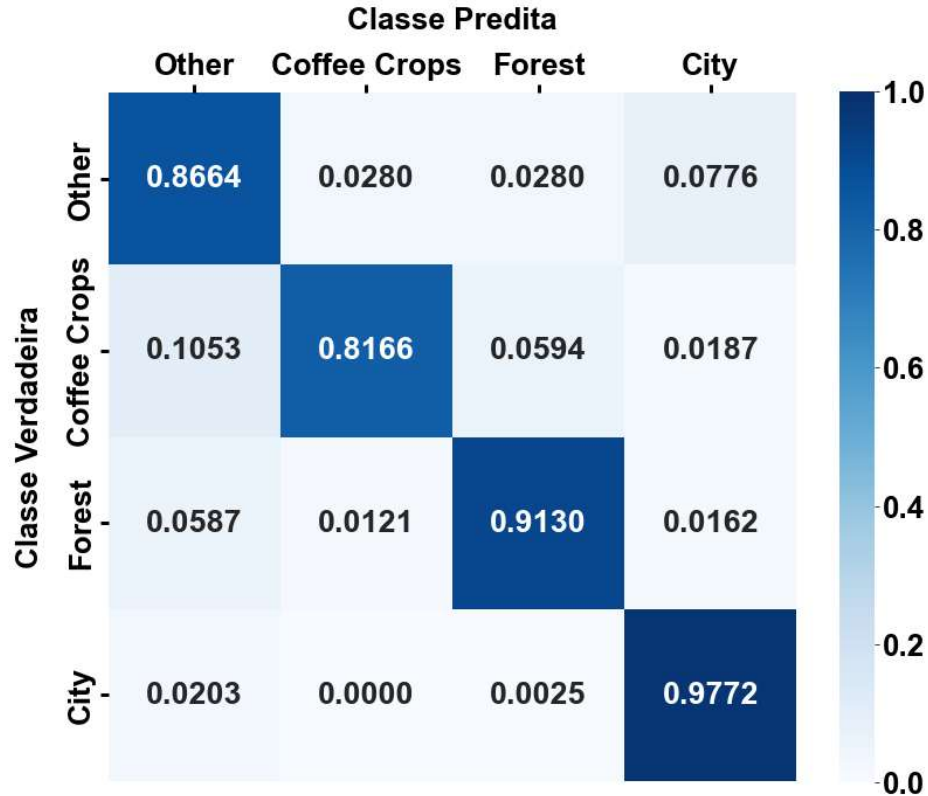
Métricas	Other	Coffee Crops	Forest	City
F1-score	0,8450	0,8796	0,9117	0,9353
IoU	0,7316	0,7851	0,8377	0,8784

A classe City apresentou F1-score de 0,9353 e IoU de 0,8784, indicando que a classe City foi identificada com alta precisão devido às suas características espectrais distintas. A classe Forest obteve F1-score de 0,9117 e IoU de 0,8377, evidenciando a capacidade do modelo em distinguir as áreas florestais.

Para a classe de interesse principal, Coffee Crops, o modelo alcançou um F1-score de 0,8796 e IoU de 0,7851, uma redução em relação ao desempenho no conjunto de validação, demonstrando a complexidade espectral desta classe e sua similaridade com outras coberturas vegetais. A classe Other, que engloba diversas coberturas não classificadas nas demais categorias, apresentou F1-score de 0,8450 e IoU de 0,7316. Este resultado é esperado devido à heterogeneidade desta classe, que agrupa diferentes tipos de cobertura do solo com características espectrais variadas.

Complementando a análise estatística, foi gerada uma matriz de confusão a partir da aplicação do modelo com a combinação de bandas (G, R, NIR) nas cinco cenas de teste, ilustrada na Figura 6. Esta matriz permite uma compreensão mais detalhada dos padrões de classificação e dos tipos específicos de erros cometidos pelo modelo nas novas cenas.

Figura 6. Matriz de confusão obtida a partir da segmentação das cinco cenas de teste pelo modelo com a combinação de bandas (G, R, NIR).



Fonte: Do Autor

Os resultados da matriz de confusão corroboram as métricas estatísticas obtidas, indicando que o modelo manteve seu desempenho diante de dados não vistos durante o treinamento e validação. A classe City apresentou a maior taxa de acerto (0,9772), confirmando os altos valores de F1-score (0,9353) e IoU (0,8784) observados na análise por classe. Similarmente, a classe Forest demonstrou alta precisão com taxa de acerto de 0,9130, consistente com seu F1-score de 0,9117.

A classe Coffee Crops obteve uma taxa de acerto de 0,8166, sendo mais uma vez a categoria com maior índice de confusão, especialmente em relação à classe Other (0,1053), padrão já observado com as combinações de bandas no conjunto de validação, e que explica este desempenho desta classe nas métricas estatísticas (F1-score de 0,8796 e IoU de 0,7851). Também houve pequena confusão com Forest (0,0594), o que reforça a sobreposição espectral entre culturas cafeeiras e áreas florestais.

A classe Other apresentou taxa de acerto de 0,8664, valor condizente com seu menor desempenho nas métricas estatísticas (F1-score de 0,8450 e IoU de 0,7316). Os baixos índices de erro nas demais classes, evidenciados tanto pelas métricas quanto pela matriz de confusão,

indicam que o modelo conseguiu generalizar seu aprendizado, mantendo desempenho estável mesmo fora do conjunto de treinamento e validação.

5 DISCUSSÃO

5.1 DESEMPENHO GLOBAL DAS COMBINAÇÕES ESPECTRAIS NA SEGMENTAÇÃO

A combinação de diferentes bandas espectrais de imagens de sensoriamento remoto pode ter impacto direto na segmentação de classes de cobertura do solo. Neste estudo, isso se confirmou na avaliação do desempenho geral da rede U-Net, que variou conforme a configuração espectral adotada. Entre as configurações avaliadas, as combinações (B, G, R) e (G, R, NIR) demonstraram superioridade consistente em todas as métricas avaliadas. Indicando que essas configurações oferecem informações espectrais que facilitam a discriminação entre as diferentes classes de interesse.

A combinação (B, G, R) alcançou o melhor desempenho global, com acurácia de 0,9659 e IoU de 0,9354, evidenciando que a segmentação pode ser altamente eficaz utilizando exclusivamente bandas do espectro visível. Resultados semelhantes também foram obtidos por Ding *et al.* (2023), que observaram alto desempenho na segmentação de cobertura vegetal usando a combinação de bandas (B, G, R). Já a combinação (G, R, NIR) apresentou o segundo melhor desempenho geral (acurácia de 0,9622 e IoU de 0,9287), demonstrando que a substituição da banda azul pela banda do infravermelho próximo pode ser vantajosa em determinados contextos. Esta configuração se mostrou particularmente eficaz para a identificação da classe de interesse Coffee Crops, obtendo o maior F1-score (0,9616) e IoU (0,9260) para esta categoria específica, o que sugere que a banda NIR, em substituição da banda B, pode possuir informações espectrais relevantes para a discriminação de cultivos de café.

Para testar a capacidade de generalização do modelo e garantir que seu desempenho não estava restrito aos dados de treinamento e validação, o modelo com a combinação (G, R, NIR), que se destacou na identificação da classe de interesse, foi aplicado em cinco cenas completamente independentes. Os resultados demonstraram a robustez do modelo, que alcançou uma acurácia global de 0,8933 e um IoU de 0,8082, e alcançou bom desempenho para a classe Coffee Crops, com F1-score de 0,8796 e IoU de 0,7851 nas cenas de teste. Embora se observe uma leve e esperada queda nas métricas em comparação com o conjunto de validação,

os valores obtidos confirmam que o modelo foi capaz de generalizar para novas áreas, validando a abordagem para aplicações de mapeamento.

Por outro lado, as combinações (B, G, NIR) e especialmente (B, R, NIR) apresentaram desempenho inferior. A combinação (B, R, NIR) obteve os menores valores em todas as métricas (acurácia de 0,8935 e IoU de 0,8180), sendo particularmente deficiente na identificação da classe Coffee Crops (F1-score de 0,8114 e IoU de 0,6826). Esses resultados indicam que as combinações das bandas influenciam diretamente na capacidade do modelo de capturar características espectrais das classes com maior fidelidade, bem como indica que nem toda adição do NIR resulta em ganho de performance, ao contrário, certas combinações podem gerar redundância ou baixa complementariedade entre as bandas, prejudicando a qualidade da segmentação.

Esse impacto das combinações espectrais na capacidade de discriminar as classes também é reforçado pelas matrizes de confusão, que revelam um padrão consistente de erros na diferenciação entre Coffee Crops e Other, independentemente da combinação de bandas utilizada. Essa confusão sistemática indica uma sobreposição nas assinaturas espectrais dessas categorias, o que dificulta a classificação entre elas. Tal padrão ressalta a importância de considerar não apenas a resposta espectral individual de cada banda, mas também a complementariedade e a sinergia entre elas na configuração do modelo.

Os resultados obtidos demonstram que a inclusão da banda NIR não garante melhoria no desempenho da segmentação. A eficácia de uma combinação espectral depende da interação entre as bandas selecionadas e sua capacidade de fornecer informações complementares para a discriminação das classes de interesse. Este achado enfatiza a necessidade da seleção de bandas espectrais, considerando as características específicas do problema de segmentação e das classes envolvidas, para maximizar o desempenho em aplicações de sensoriamento remoto multiespectral.

5.2 INFLUÊNCIA DAS BANDAS ESPECTRAIS NA SEGMENTAÇÃO POR CLASSE

A análise por classe reforça a superioridade das combinações (B, G, R) e (G, R, NIR), que apresentaram os maiores valores de F1-score e IoU em praticamente todas as categorias. Já ao analisar a classe Coffee Crops, as combinações (B, G, R) e (G, R, NIR) foram as únicas que mantiveram níveis elevados de F1-score 0,9599 e 0,9616, e IoU 0,9228 e 0,9260, respectivamente, demonstrando maior eficácia na segmentação dessa categoria. Esses

resultados indicam que tanto o espectro visível completo quanto a introdução do NIR em substituição à banda B são eficientes na segmentação da classe Coffee Crops. A presença das bandas G e R em ambas as combinações parece ser essencial, provavelmente por capturarem informações importantes relacionadas ao padrão espectral característico do café.

Observa-se que, ao remover a banda R ou G nas combinações, o desempenho na segmentação cai, principalmente com a remoção da banda G. Isso pode ocorrer devido à grande importância dessa banda para diferenciar a classe Coffee Crops, indicando que ela está intimamente relacionada à discriminação em relação às outras classes. Essa observação vai ao encontro dos achados de Motohka *et al.* (2010), que demonstraram a relevância da relação entre as bandas G e R, além de destacar a maior contribuição da banda G na distinção entre diferentes tipos de vegetação.

Esse padrão de queda de desempenho torna-se ainda mais evidente ao se observar os resultados das combinações (B, G, NIR) e (B, R, NIR) que apresentaram performance notavelmente inferior para a classe Coffee Crops. Os valores de F1-score para essa classe caíram para 0,9422 e 0,8114, respectivamente, com IoU de 0,8907 e 0,6826, indicando menor eficácia na segmentação dessa classe. A ausência da banda R em (B, G, NIR) e da banda G em (B, R, NIR) compromete a capacidade do modelo de capturar os padrões espectrais mais relevantes para diferenciar as lavouras de café de outras formas de vegetação.

Esse comportamento pode ser atribuído às características morfofisiológicas do cafeeiro, especialmente à alta concentração de clorofila presente em suas folhas, que é significativamente diferente da encontrada em outros tipos de vegetação, conforme demonstrado por Falcioni *et al.* (2023). Essa composição específica provoca um padrão de reflectância característico da cultura, com forte absorção na faixa do R e alta reflectância na faixa do G, criando um contraste espectral que favorece a distinção da cultura em relação a outras vegetações. A complementaridade entre as bandas G e R se mostra, portanto, essencial para realçar essas diferenças, explicando a queda de desempenho quando uma delas é removida.

Por outro lado, a banda B é menos específica, pois diversos pigmentos, como clorofilas e carotenoides a absorvem, o que reduz seu poder discriminativo. Assim, as bandas G e R capturam de forma mais eficaz o comportamento espectral característico das lavouras de café, sendo determinantes para a segmentação dessa classe. A perda de qualquer uma dessas bandas compromete diretamente a capacidade do modelo de distinguir as lavouras de café de outras coberturas vegetais. Dessa forma, os resultados sugerem que a segmentação por classe exige atenção especial à composição espectral, pois combinações inadequadas podem comprometer

significativamente a capacidade do modelo em identificar categorias específicas, mesmo quando o desempenho geral parece satisfatório.

5.3 INTERPRETAÇÃO DA CONFUSÃO ESPECTRAL ENTRE CLASSES

A análise dos resultados revela um padrão de discrepância entre as métricas F1-score e IoU para a classe Coffee Crops. Enquanto o F1-score, que representa a média harmônica entre precisão e recall, manteve valores mais elevados, os valores correspondentes de IoU, que penaliza mais severamente erros de predição, foram consistentemente inferiores. Essa diferença, observada em todas as combinações de bandas no conjunto de validação, também se manifestou claramente nos testes com as cinco cenas independentes.

A principal causa para essa discrepância é a confusão sistemática entre as classes Coffee Crops e Other. A análise da matriz de confusão do conjunto de validação (Figura 3) já apontava a sobreposição espectral entre essas classes como o principal fator de erro, conforme as combinações de bandas: 0,0707 para (B, G, R), 0,0799 para (B, G, NIR), chegando a 0,3128 para (B, R, NIR), e 0,0651 para (G, R, NIR). Esse padrão também é encontrado nos resultados nas cinco cenas de teste, onde a matriz de confusão gerada para esses novos dados (Figura 6) revela que o mesmo padrão de erro não apenas persistiu, mas continuou sendo o mais significativo: 10,53% dos pixels de Coffee Crops foram classificados incorretamente como Other, representando um aumento em relação aos 6,51% observados no conjunto de validação com a combinação (G, R, NIR).

Adicionalmente, a análise em novas áreas revelou uma confusão adicional da classe Coffee Crops com a classe Forest (0,0594), que não era tão evidente no conjunto de validação. Esta nova fonte de erro indica que, em cenários reais de aplicação, as plantações de café podem ser confundidas não apenas com outras classes heterogêneas (Other), mas também com áreas florestais, especialmente em regiões onde o café é cultivado próximo a fragmentos florestais. Este achado é particularmente relevante para a região das Matas de Minas, caracterizada pela presença de transição entre cultivos de café e áreas florestais.

Essa sobreposição espectral explica por que o modelo mantém bons valores de F1-score (0,8796) mas apresenta IoU inferior (0,7851) para Coffee Crops nas cenas de teste. O modelo demonstra capacidade de identificar adequadamente a presença de culturas cafeeiras, o que resulta em altos valores de F1-score, mas enfrenta dificuldades na delimitação precisa dos contornos espaciais dessas áreas, especialmente nas bordas onde ocorre a transição entre Coffee

Crops e Other. Esta limitação é mais severamente penalizada pelo IoU, que considera tanto a sobreposição quanto a união entre as áreas preditas e reais. A escolha da combinação de bandas também influencia diretamente essa divergência, a combinação (B, R, NIR) apresentou a maior discrepância entre F1-score (0,8114) e IoU (0,6826), enquanto as combinações (B, G, R) e (G, R, NIR) mostraram melhor alinhamento.

A redução do desempenho nas cenas de teste, de IoU 0,9260 no conjunto de validação com a combinação de bandas (G, R, NIR) para 0,7851, confirma que a segmentação da classe Coffee Crops é mais desafiadora em cenários reais do que em dados controlados. Esta diferença ressalta a importância da avaliação em múltiplas cenas independentes para uma compreensão realista das limitações do modelo. Além disso, sugere que estratégias específicas, como subdividir a classe Other em mais classes, subdividindo em categorias mais detalhadas, como pastagens, solos expostos e culturas específicas, como milho e cana-de-açúcar, podem ser necessárias para melhorar a robustez da segmentação em aplicações práticas.

Portanto, a diferença observada entre F1-score e IoU, tanto no conjunto de validação quanto nas cenas de teste, não indica uma falha do modelo em detectar a presença da cultura, mas sim evidencia os desafios inerentes à delimitação exata de classes espectralmente semelhantes. Esta interpretação é crucial para compreender que o modelo mantém uma boa capacidade de identificação, mas que a precisão de seus contornos pode ser aprimorada em trabalhos futuros, especialmente para distinguir a classe Coffee Crops de outras vegetações de comportamento espectral similar, como as contidas na classe Other.

6 CONCLUSÃO

Neste estudo, foi explorada a aplicabilidade da rede de segmentação semântica U-Net para a segmentação de lavouras de café na região das Matas de Minas, utilizando imagens multiespectrais do PlanetScope. A pesquisa evidenciou que a escolha adequada da combinação de bandas espectrais é um fator determinante para o desempenho da segmentação, influenciando diretamente a capacidade do modelo de discriminar as culturas cafeeiras de outras classes de cobertura do solo.

O estudo demonstrou que as combinações espectrais (B, G, R) e (G, R, NIR) possuem desempenho superior geral na segmentação. Especificamente para a classe Coffee Crops, a combinação (G, R, NIR) demonstrou a melhor performance. E esta combinação demonstrou

capacidade de generalização ao ser aplicado em cinco cenas independentes, mantendo desempenho na segmentação, confirmando que pode ser aplicada em novas áreas.

A análise revelou que a presença simultânea das bandas G e R é fundamental para a segmentação de lavouras de café, provavelmente devido às características morfofisiológicas específicas do cafeeiro. A remoção de qualquer uma dessas bandas resultou em queda do desempenho, como observado nas combinações (B, G, NIR) e (B, R, NIR).

A principal limitação identificada foi a confusão sistemática entre as classes Coffee Crops e Other, revelada pelas matrizes de confusão, que indica sobreposição espectral entre essas categorias. Esta confusão foi responsável pela discrepância observada entre os valores de F1-score e IoU, especialmente para a classe de interesse, sugerindo que embora o modelo identifique adequadamente a presença de culturas cafeeiras, enfrenta desafios na delimitação espacial precisa dessas áreas, particularmente nas regiões de transição entre classes.

REFERÊNCIAS

ANDERSON, Helen B. *et al.* Using Ordinary Digital Cameras in Place of Near-Infrared Sensors to Derive Vegetation Indices for Phenology Studies of High Arctic Vegetation. *Remote Sensing*, v. 8, n. 10, p. 847, 17 out. 2016.

AYHAN, Bulent *et al.* Vegetation Detection Using Deep Learning and Conventional Methods. *Remote Sensing*, v. 12, n. 15, p. 2502, 4 ago. 2020.

BAETA, Rafael *et al.* Learning Deep Features on Multiple Scales for Coffee Crop Recognition. *Proceedings - 30th Conference on Graphics, Patterns and Images, SIBGRAPI 2017*, p. 262–268, 3 nov. 2017.

BHUIYAN, Md Abul Ehsan *et al.* Understanding the Effects of Optimal Combination of Spectral Bands on Deep Learning Model Predictions: A Case Study Based on Permafrost Tundra Landform Mapping Using High Resolution Multispectral Satellite Imagery. *Journal of Imaging*, v. 6, n. 9, p. 97, 17 set. 2020.

BOURGOIN, Clément *et al.* Assessing the ecological vulnerability of forest landscape to agricultural frontier expansion in the Central Highlands of Vietnam. *International Journal of Applied Earth Observation and Geoinformation*, v. 84, p. 101958, 1 fev. 2020a.

BOURGOIN, Clément *et al.* Assessing the ecological vulnerability of forest landscape to agricultural frontier expansion in the Central Highlands of Vietnam. *International Journal of Applied Earth Observation and Geoinformation*, v. 84, p. 101958, 1 fev. 2020b.

BUTLER, Keith T. *et al.* Machine learning for molecular and materials science. *Nature* 2018 559:7715, v. 559, n. 7715, p. 547–555, 25 jul. 2018.

CHANG, Hye Yoon *et al.* Artificial Intelligence in Pathology. *Journal of Pathology and Translational Medicine*, v. 53, n. 1, p. 1–12, 28 dez. 2018.

CHANG, Zhu *et al.* Crop Type Identification Using High-Resolution Remote Sensing Images Based on an Improved DeepLabV3+ Network. *Remote Sensing*, v. 15, n. 21, p. 5088, 24 out. 2023.

CHEMURA, Abel; MUTANGA, Onesimo. Developing detailed age-specific thematic maps for coffee (*Coffea arabica* L.) in heterogeneous agricultural landscapes using random forests applied on Landsat 8 multispectral sensor. *Geocarto International*, v. 32, n. 7, p. 759–776, 3 jul. 2017.

CORREIA, Paula Ferreira da Cruz *et al.* Impacts of Brazilian Green Coffee Production and Its Logistical Corridors on the International Coffee Market. *Logistics 2024*, Vol. 8, Page 39, v. 8, n. 2, p. 39, 9 abr. 2024.

CREWELL, Susanne; WENDISCH, Manfred; LÖHNERT, Ulrich. Passive Solar and Microwave Spectral Radiometers. *Springer Handbooks*, p. 821–840, 2021.

DAMATTA, Fábio M. *et al.* Why could the coffee crop endure climate change and global warming to a greater extent than previously estimated? *Climatic Change*, v. 152, n. 1, p. 167–178, 30 jan. 2019.

DE CARVALHO ALVES, Marcelo *et al.* Multisensor analysis for environmental targets identification in the region of Funil dam, state of Minas Gerais, Brazil. *Applied Geomatics*, v. 15, n. 4, p. 807–827, 1 dez. 2023.

DEFRIES, Ruth. Remote Sensing and Image Processing. *Encyclopedia of Biodiversity: Second Edition*, p. 389–399, 1 jan. 2013.

DING, Binbin *et al.* Land Cover Extraction in the Typical Black Soil Region of Northeast China Using High-Resolution Remote Sensing Imagery. *Land*, v. 12, n. 8, p. 1566, 7 ago. 2023.

DUAN, Si Bo *et al.* Land Surface Temperature Retrieval from Passive Microwave Satellite Observations: State-of-the-Art and Future Directions. *Remote Sensing*, v. 12, n. 16, p. 2573, 10 ago. 2020.

EDWARDS, Audrene S.; KAPLAN, Bruce; JIE, Tun. A Primer on Machine Learning. *Transplantation*, v. 105, n. 4, p. 699–703, 1 abr. 2021.

EMBURY, Owen *et al.* Satellite-based time-series of sea-surface temperature since 1980 for climate applications. *Nature*, v. 11, n. 1, p. 1–22, 29 mar. 2024.

FALCIONI, Renan *et al.* Reflectance Spectroscopy for the Classification and Prediction of Pigments in Agronomic Crops. *Plants*, v. 12, n. 12, p. 2347, 1 jun. 2023.

FARIA, Luiz Henrique Portela; BARREIRA, Luiz Carlos. Between tenements and coffee: the construction of modernity in the city of Santos (1870-1920). *STUDIES IN MULTIDISCIPLINARY REVIEW*, v. 6, n. 1, p. e15684, 22 abr. 2025.

FERREIRA, Thiago *et al.* Introduction to Coffee Plant and Genetics. *Coffee*, p. 1–25, 11 jan. 2019.

GAERTNER, Julie *et al.* Vegetation classification of Coffea on Hawaii Island using WorldView-2 satellite imagery. *Journal of Applied Remote Sensing*, v. 11, n. 4, p. 046005, 13 out. 2017.

GREENER, Joe G. *et al.* A guide to machine learning for biologists. *Nature Reviews Molecular Cell Biology*, v. 23, n. 1, p. 40–55, 1 jan. 2022.

HAN, Jiawei; PEI, Jian; TONG, H. *Data mining: concepts and techniques*. 4th. ed. [S.l.]: Morgan Kaufmann, 2022.

HATFIELD, Jerry L. *et al.* Applications of Vegetative Indices from Remote Sensing to Agriculture: Past and Future. *Inventions*, v. 4, n. 4, p. 71, 6 dez. 2019.

HINTON, G. E.; SALAKHUTDINOV, R. R. Reducing the dimensionality of data with neural networks. *Science*, v. 313, n. 5786, p. 504–507, 28 jul. 2006.

HUYNH, Tri; NIBALI, Aiden; HE, Zhen. Semi-supervised learning for medical image classification using imbalanced training data. *Computer Methods and Programs in Biomedicine*, v. 216, p. 106628, 1 abr. 2022.

INPI. FICHA TÉCNICA DE REGISTRO DE INDICAÇÃO GEOGRÁFICA. [S.l.: S.n.]. Disponível em: <http://revistas.inpi.gov.br/pdf/Indicacoes_Geograficas2606.pdf>.

INPI. FICHA TÉCNICA DE REGISTRO DE INDICAÇÃO GEOGRÁFICA. INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL, 15 dez. 2020b.

INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL. Indicações Geográficas. Disponível em: <<https://www.gov.br/inpi/pt-br/acesso-a-informacao/perguntas-frequentes/indicacoes-geograficas>>. Acesso em: 11 set. 2025.

INTERNATIONAL COFFEE ORGANIZATION. THE FUTURE OF COFFEE Investing in youth for a resilient and sustainable coffee sector. INTERNATIONAL COFFEE ORGANIZATION, 2021.

IQBAL, Iram M. *et al.* Identifying the spectral signatures of invasive and native plant species in two protected areas of pakistan through field spectroscopy. *Remote Sensing*, v. 13, n. 19, p. 4009, 1 out. 2021.

ISLAM, T. *et al.* Remote sensing of aerosols, clouds, and precipitation. [S.l.]: Elsevier, 2017.

JAVOID, Mohd *et al.* Sensors for daily life: A review. *Sensors International*, v. 2, p. 100121, 1 jan. 2021.

JIA, Peiyan *et al.* Semantic segmentation of deep learning remote sensing images based on band combination principle: Application in urban planning and land use. *Computer Communications*, v. 217, p. 97–106, 1 mar. 2024.

KOUADIO, Louis *et al.* Probabilistic yield forecasting of robusta coffee at the farm scale using agroclimatic and remote sensing derived indices. *Agricultural and Forest Meteorology*, v. 306, p. 108449, 15 ago. 2021.

LECUN, Yann; BENGIO, Yoshua; HINTON, Geoffrey. Deep learning. *Nature*, v. 521, n. 7553, p. 436–444, 27 maio 2015.

LELONG, Camille C. D.; THONG-CHANE, Audrey. Application of textural analysis on very high resolution panchromatic images to map coffee orchards in Uganda. *International Geoscience and Remote Sensing Symposium (IGARSS)*, v. 2, p. 1007–1009, 2003.

LONG, Jonathan; SHELHAMER, Evan; DARRELL, Trevor. Fully Convolutional Networks for Semantic Segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition, 2015.* . Acesso em: 9 maio. 2025

MANJUNATH, K. R.; RAY, Shibendu Shankar; PANIGRAHY, Sushma. Discrimination of Spectrally-Close Crops Using Ground-Based Hyperspectral Data. *Journal of the Indian Society of Remote Sensing*, v. 39, n. 4, p. 599–602, 3 dez. 2011.

MARBLE, D. F.; PEUQUET, D. J. Geographic Information Systems and Remote Sensing. *In: Manual of remote sensing, 2nd edition. Vol. I. [S.l.]*: Chandos Publishing, 2017. p. 71–85.

MARTIN, Randall V. Satellite remote sensing of surface air quality. *Atmospheric Environment*, v. 42, n. 34, p. 7823–7843, 1 nov. 2008.

MARTINEZ, Herminia Emilia Prieto *et al.* Agronomic practices toward coffee sustainability. A review. *Scientia Agricola*, v. 81, p. e20220277, 2024.

MARTÍNEZ-VERDUZCO, Guillermo C.; GALEANA-PIZAÑA, J. Mauricio; CRUZ-BELLO, Gustavo M. Coupling Community Mapping and supervised classification to discriminate Shade coffee from Natural vegetation. *Applied Geography*, v. 34, p. 1–9, 1 maio 2012.

MATSUO, Yutaka *et al.* Deep learning, reinforcement learning, and world models. *Neural Networks*, v. 152, p. 267–275, 1 ago. 2022.

MCCULLOCH, Warren S.; PITTS, Walter. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, v. 5, n. 4, p. 115–133, dez. 1943.

MO, Yujian *et al.* Review the state-of-the-art technologies of semantic segmentation based on deep learning. *Neurocomputing*, v. 493, p. 626–646, 7 jul. 2022.

MOREIRA, Eder Paulo *et al.* TOPOGRAPHIC EFFECT ON SPECTRAL VEGETATION INDICES FROM LANDSAT TM DATA: IS TOPOGRAPHIC CORRECTION NECESSARY? *Boletim de Ciências Geodésicas*, v. 22, n. 1, p. 95–107, 1 jan. 2016.

MOREIRA, Maurício Alves; BARROS, Marco Aurélio; RUDORFF, Bernardo Friedrich Theodor. Geotechnologies in Coffee crop mapping at municipality scale. *Sociedade & Natureza*, v. 20, n. 1, p. 101–110, jun. 2008.

MOTOHKA, Takeshi *et al.* Applicability of Green-Red Vegetation Index for Remote Sensing of Vegetation Phenology. *Remote Sensing*, v. 2, n. 10, p. 2369–2387, 15 out. 2010.

MUTHUKRISHNAN, Nikesh *et al.* Brief History of Artificial. Machine Learning and Other Artificial Intelligence Applications, An Issue of Neuroimaging Clinics of North America, E-Book: Machine Learning and Other Artificial Intelligence Applications, An Issue of Neuroimaging Clinics of North America, v. 30, n. 4, p. 393, 2020.

NOVO, EMLM. *Sensoriamento Remoto: Princípios e Aplicações*. 2ª. Edição. Edgar Blucher, São Paulo. 308p, 1992.

OTUKEI, J. R.; BLASCHKE, T. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. *International Journal of Applied Earth Observation and Geoinformation*, v. 12, n. SUPPL. 1, p. S27–S31, 1 fev. 2010.

PANCSIRA, János. International Coffee Trade: a literature review. *Journal of Agricultural informatics*, v. 13, n. 1, 2022.

PICOLI, M. C. A. *et al.* UNSUPERVISED SEGMENTATION OF SMALLHOLDER FIELDS IN MOZAMBIQUE USING PLANETSCOPE IMAGERY. 2020.

QUEIROZ, DM de *et al.* *Digital Agriculture*. [S.l.]: Springer, 2022.

RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 9351, p. 234–241, 2015a.

RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 9351, p. 234–241, 2015b.

ROY, David P. *et al.* A global analysis of the temporal availability of PlanetScope high spatial resolution multi-spectral imagery. *Remote Sensing of Environment*, v. 264, p. 112586, 1 out. 2021.

SARKER, Iqbal H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, v. 2, n. 3, p. 1–21, 1 maio 2021.

SCHMITT-HARSH, Mikaela. Landscape change in Guatemala: Driving forces of forest and coffee agroforest expansion and contraction from 1990 to 2010. *Applied Geography*, v. 40, p. 40–50, 1 jun. 2013a.

SCHMITT-HARSH, Mikaela. Landscape change in Guatemala: Driving forces of forest and coffee agroforest expansion and contraction from 1990 to 2010. *Applied Geography*, v. 40, p. 40–50, 1 jun. 2013b.

SHINDE, Pramila P.; SHAH, Seema. A Review of Machine Learning and Deep Learning Applications. Proceedings - 2018 4th International Conference on Computing, Communication Control and Automation, ICCUBEA 2018, 2 jul. 2018.

SHRESTHA, Ajay; MAHMOOD, Ausif. Review of deep learning algorithms and architectures. IEEE Access, v. 7, p. 53040–53065, 2019.

SILVA, Wagner F. *et al.* Discrimination of agricultural crops in a tropical semi-arid region of Brazil based on L-band polarimetric airborne SAR data. ISPRS Journal of Photogrammetry and Remote Sensing, v. 64, n. 5, p. 458–463, 1 set. 2009a.

SILVA, Wagner F. *et al.* Discrimination of agricultural crops in a tropical semi-arid region of Brazil based on L-band polarimetric airborne SAR data. ISPRS Journal of Photogrammetry and Remote Sensing, v. 64, n. 5, p. 458–463, 1 set. 2009b.

SOTT, Michele Kremer *et al.* Precision Techniques and Agriculture 4.0 Technologies to Promote Sustainability in the Coffee Sector: State of the Art, Challenges and Future Trends. IEEE Access, v. 8, p. 149854–149867, 2020.

SOVANN, Chansopheaktra *et al.* Importance of Spectral Information, Seasonality, and Topography on Land Cover Classification of Tropical Land Cover Mapping. Remote Sensing, v. 17, n. 9, p. 1551, 1 maio 2025.

SPENCE, Charles; CARVALHO, Fabiana M. The coffee drinking experience: Product extrinsic (atmospheric) influences on taste and choice. Food Quality and Preference, v. 80, p. 103802, 1 mar. 2020.

TAPAKIS, R.; CHARALAMBIDES, A. G. Equipment and methodologies for cloud detection and classification: A review. Solar Energy, v. 95, p. 392–430, 1 set. 2013.

TONG, Xin Yi; XIA, Gui Song; ZHU, Xiao Xiang. Enabling country-scale land cover mapping with meter-resolution satellite imagery. ISPRS Journal of Photogrammetry and Remote Sensing, v. 196, p. 178–196, 1 fev. 2023.

TOPIK, Steven. The World Coffee Market in the Eighteenth And Nineteenth Centuries, from Colonial To National Regimes. 2004.

TOPIK, Steven. How Brazil Expanded the World Coffee Economy. *Österreichische Zeitschrift für Geschichtswissenschaften*, v. 30, n. 3, p. 11–41, 16 dez. 2019.

UTRILLA-CATALAN, Rebeca *et al.* Growing Inequality in the Coffee Global Value Chain: A Complex Network Assessment. *Sustainability* 2022, Vol. 14, Page 672, v. 14, n. 2, p. 672, 8 jan. 2022.

VALÉRIE, Poncet *et al.* Which diversification trajectories make coffee farming more sustainable? *Current Opinion in Environmental Sustainability*, v. 68, p. 101432, 1 jun. 2024.

VAN DER VOSSSEN, Herbert; BERTRAND, Benoît; CHARRIER, André. Next generation variety development for sustainable production of arabica coffee (*Coffea arabica* L.): a review. *Euphytica*, v. 204, n. 2, p. 243–256, 19 jul. 2015.

VERBURG, René *et al.* An innovation perspective to climate change adaptation in coffee systems. *Environmental Science & Policy*, v. 97, p. 16–24, 1 jul. 2019.

WANG, Hanzhao *et al.* SegForest: A Segmentation Model for Remote Sensing Images. *Forests*, v. 14, n. 7, p. 1509, 24 jul. 2023.

WANG, Jun *et al.* Integration of Remote Sensing and Machine Learning for Precision Agriculture: A Comprehensive Perspective on Applications. *Agronomy* 2024, Vol. 14, Page 1975, v. 14, n. 9, p. 1975, 1 set. 2024.

WINKER, David *et al.* Observational Constraints on Cloud Feedbacks: The Role of Active Satellite Sensors. *Surveys in Geophysics*, v. 38, n. 6, p. 1483–1508, 1 nov. 2017.

YASAKA, Koichiro *et al.* Deep learning with convolutional neural network in radiology. *Japanese Journal of Radiology*, v. 36, n. 4, p. 257–272, 1 abr. 2018.

ZHANG, J.; LI, J. Spatial Cognitive Engine Technology. [S.l.: S.n.].

ZHANG, Xin *et al.* How Well Do Deep Learning-Based Methods for Land Cover Classification and Object Detection Perform on High Resolution Remote Sensing Imagery? *Remote Sensing*, v. 12, n. 3, p. 417, 28 jan. 2020.

ZHOU, Lei *et al.* Machine learning for combustion. *Energy and AI*, v. 7, p. 100128, 1 jan. 2022.

ZUO, C. *et al.* Deep learning in optical metrology: a review. *Nature*, v. 11, n. 1, p. 1–54, 2022.