



Universidade Federal do Piauí
Centro de Ciências da Natureza
Programa de Pós-Graduação em Ciência da Computação

Uma Abordagem Híbrida de CNN para Classificação e Mapeamento Hiperespectral de Áreas Agrícolas

Robert Silva Lima

Teresina-PI, Março de 2025

Robert Silva Lima

Uma Abordagem Híbrida de CNN para Classificação e Mapeamento Hiperespectral de Áreas Agrícolas

Defesa de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da UFPI (área de concentração: Visão Computacional), como parte dos requisitos necessários para a obtenção do Título de Mestre em Ciência da Computação.

Universidade Federal do Piauí – UFPI

Centro de Ciências da Natureza

Programa de Pós-Graduação em Ciência da Computação

Orientador: Prof. Dr. Romuere Rodrigues Veloso e Silva

Teresina-PI

Março de 2025

FICHA CATALOGRÁFICA
Universidade Federal do Piauí
Sistema de Bibliotecas UFPI - SIBi/UFPI
Biblioteca Setorial do CCN

L732a Lima, Robert Silva.
 Uma abordagem híbrida de CNN para classificação e
 mapeamento hiperespectral de áreas agrícolas / Robert Silva
 Lima. -- 2024.
 95 f.

 Dissertação (Mestrado) - Universidade Federal do Piauí.
 Centro de Ciências da Natureza. Programa de Pós-Graduação
 em Ciências da Computação, Teresina, 2024.
 “Orientador: Prof. Dr. Romuere Rodrigues Veloso e Silva”.

 1. Sensoriamento Remoto hiperespectral. 2. Aprendizado
 de máquina. 3. Redes Neurais Convolucionais - CNN. 4.
 Mapeamento agrícola. I. Silva, Romuere Rodrigues Veloso e.
 II. Título.

CDD 005.2

Bibliotecária: Caryne Maria da Silva Gomes - CRB3/1461

Robert Silva Lima

Uma Abordagem Híbrida de CNN para Classificação e Mapeamento Hiperespectral de Áreas Agrícolas

Defesa de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da UFPI (área de concentração: Visão Computacional), como parte dos requisitos necessários para a obtenção do Título de Mestre em Ciência da Computação.

Trabalho aprovado. Teresina-PI, 10 de Março de 2025:

Documento assinado digitalmente
 **ROMUERE RODRIGUES VELOSO E SILVA**
Data: 19/03/2025 16:06:48-0300
Verifique em <https://validar.iti.gov.br>

Prof. Romuere Rodrigues Veloso e Silva
(Presidente da banca examinadora)

Documento assinado digitalmente
 **FLAVIO HENRIQUE DUARTE DE ARAUJO**
Data: 19/03/2025 16:01:06-0300
Verifique em <https://validar.iti.gov.br>

Prof. Flávio Henrique Duarte de Araujo
(Examinador externo ao programa)

Documento assinado digitalmente
 **IALIS CAVALCANTE DE PAULA JÚNIOR**
Data: 19/03/2025 15:17:03-0300
Verifique em <https://validar.iti.gov.br>

Prof. Iális Cavalcante de Paula Júnior
(Examinador externo à instituição)

Documento assinado digitalmente
 **VINICIUS PONTE MACHADO**
Data: 19/03/2025 15:36:46-0300
Verifique em <https://validar.iti.gov.br>

Prof. Vinícius Ponte Machado
(Examinador interno)

Teresina-PI
Março de 2025

*A minha família, minha esposa Bruna Furtado Silva Lima e minhas filhas Júlia Roberta
Silva Lima e Giovana Graziela Silva Lima por estarem sempre ao meu lado.*

Agradecimentos

Agradeço primeiro a Deus por tornar esse sonho possível.

Agradeço à minha família, minha esposa Bruna Furtado Silva Lima e minhas filhas Júlia Roberta Silva Lima e Giovana Graziela Silva Lima por estarem sempre ao meu lado.

Agradeço aos meus pais, Zilmar Silva Lima e Raimundo Barroso Lima, pela educação ministrada.

Agradeço ao meu orientador, Prof. Dr. Romuere Rodrigues Veloso e Silva, e meu co-orientador Prof. Dr. Bruno Vicente Alves de Lima, por todos os conselhos, pela paciência e ajuda nesse período.

Agradeço ao meu amigo Francisco Silva Barbosa pelo auxílio em minha jornada.

Aos professores que compartilharam seus conhecimentos nas disciplinas ministradas, meus sinceros agradecimentos a todos que contribuíram para a realização deste estudo.

Ao Instituto Federal do Maranhão, por permitir meu afastamento para realizar meu mestrado.

E, por fim, àqueles presentes à defesa de Mestrado.

*“Tudo posso,
naquele que me fortalece.”
(Filipenses 4:13)*

Resumo

A classificação hiperespectral de áreas agrícolas enfrenta desafios significativos devido à complexidade dos dados espectrais e à necessidade de precisão na distinção entre diferentes tipos de culturas e variações de solo. Para superar esses desafios, esta dissertação propõe uma abordagem híbrida que integra Redes Neurais Convolucionais (CNNs) para aprimorar a classificação e o mapeamento de imagens hiperespectrais em atividades agrícolas. O modelo proposto combina arquiteturas convolucionais bidimensionais (2D-CNN) e tridimensionais (3D-CNN), onde as camadas 3D-CNN processam inicialmente as imagens hiperespectrais para extrair características espectrais detalhadas, considerando a correlação entre bandas espectrais adjacentes. Em seguida, uma camada 2D-CNN captura detalhes espaciais, facilitando a distinção entre variações sutis no terreno. Como resultado, a abordagem foi testada em várias bases de dados hiperespectrais, incluindo WHU-Hi-LongKou, HanChuan, HongHu, University of Pavia, Salinas, Indian Pines, KSC e Botswana, com a quantidade de bandas variando de 103 a 274. A proposta alcançou acurácia global (OA) entre 99,90% e 100%, recall de até 100% e F1-score de até 100%, confirmando a eficácia e robustez do modelo. Este estudo contribui significativamente para o avanço na classificação hiperespectral e fornece informações importantes para o monitoramento eficiente e a gestão de recursos agrícolas. A performance do modelo indica potenciais aplicações em larga escala na agricultura.

Palavras-chaves: Sensoriamento remoto hiperespectral, Aprendizado de máquina, Aprendizado profundo, Processamento de imagens, Mapeamento agrícola.

Abstract

Hyperspectral classification of agricultural areas faces significant challenges due to the complexity of spectral data and the need for high precision in distinguishing between different crop types and soil variations. To overcome these challenges, this master's dissertation proposes a hybrid approach that integrates Convolutional Neural Networks (CNNs) to enhance the classification and mapping of hyperspectral images in agricultural activities. The model combines two-dimensional (2D-CNN) and three-dimensional (3D-CNN) convolutional architectures, where 3D-CNN layers initially process the hyperspectral images to extract detailed spectral features by considering the correlation between adjacent spectral bands. Then, a 2D-CNN layer captures spatial details, facilitating the distinction of subtle terrain variations. This combination allows the model to perform an integrated and detailed representation of the data, essential for accurate agricultural area classification. As a result, the approach was tested on various hyperspectral datasets, including WHU-Hi-LongKou, HanChuan, HongHu, University of Pavia, Salinas, Indian Pines, KSC, and Botswana, with the number of bands ranging from 103 to 274. The proposed method achieved overall accuracy (OA) between 99.90% and 100%, recall up to 100%, and F1-score up to 100%. The test loss ranged from 1.0000×10^{-7} to 1.5093×10^{-3} , along with remarkable results in metrics such as Average Accuracy (AA) and Kappa, confirming the model's effectiveness and robustness. This study contributes significantly to advances in hyperspectral classification and provides important insights for efficient monitoring and management of agricultural resources. The model's success suggests potential large-scale applications in agriculture, promoting land-use optimization and maximizing sustainable agricultural production.

Keywords: Hyperspectral remote sensing, Machine learning, Deep learning, Image processing, Agricultural mapping.

Lista de ilustrações

Figura 1 – Exemplo de captura de imagem Hiperespectral. Imagem adaptada do trabalho de Zhong et al. (2020).	6
Figura 2 – CNN. Imagem adaptada do trabalho de ((ZHONG et al., 2020))	9
Figura 3 – Representação de convolução padrão 2D e 3D. Imagem adaptada do site ((https://paperswithcode.com/method/3d-convolution))	11
Figura 4 – Verdade terrestre do conjunto de dados Indian Pines	12
Figura 5 – Verdade terrestre do conjunto de dados Salinas	14
Figura 6 – Verdade terrestre do conjunto de dados Universidade de Pavia	15
Figura 7 – Band 12 - KSC	17
Figura 8 – Band 101 - Botswana	18
Figura 9 – (a) Cubo de imagem. (b) Verdade terrestre. (c) Fotos típicas das culturas na área de estudo. Fonte: (ZHONG et al., 2020)	20
Figura 10 – (a) Cubo de imagem. (b) Verdade terrestre. (c) Fotos típicas das culturas na área de estudo. Fonte: (ZHONG et al., 2020)	21
Figura 11 – (a) Cubo de imagem. (b) Verdade terrestre. (c) Fotos típicas das culturas na área de estudo. Fonte: (ZHONG et al., 2020)	21
Figura 12 – Representação do modelo proposto neste trabalho.	41
Figura 13 – Análise de Componentes Principais (PCA) para Redução de Dimensionalidade	51
Figura 14 – Evolução da Perda e Acurácia do Modelo na Base Indian Pines	52
Figura 15 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo - Indian Pines	52
Figura 16 – Evolução da Perda e Acurácia do Modelo na Base Salinas	53
Figura 17 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo - Salinas	54
Figura 18 – Evolução da Perda e Acurácia do Modelo na Base Universidade de Pavia	55
Figura 19 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo - Base Universidade de Pavia	56
Figura 20 – Evolução da Perda e Acurácia do Modelo na Base KSC	57
Figura 21 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo - Base KSC	58
Figura 22 – Evolução da Perda e Acurácia do Modelo na Base Botswana	59
Figura 23 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo - Base de Dados Botswana	59
Figura 24 – Evolução da Perda e Acurácia do Modelo na Base LongKou	60

Figura 25 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo	
- Base LongKou	61
Figura 26 – Evolução da Perda e Acurácia do Modelo - Base HanChuan	62
Figura 27 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo	
- Base HanChuan	63
Figura 28 – Evolução da Perda e Acurácia do Modelo na Base HongHu	64
Figura 29 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo	
- Base HongHu	64

Lista de tabelas

Tabela 1 – Classes da verdade terrestre do conjunto de dados Indian Pines	13
Tabela 2 – Classes da verdade terrestre do conjunto de dados Salinas	15
Tabela 3 – Classes da verdade terrestre do conjunto de dados Universidade de Pavia	16
Tabela 4 – Classes da base de dados KSC	17
Tabela 5 – Classes da base de dados Botswana	19
Tabela 6 – Resumo dos artigos sobre Aprendizado de Máquina	27
Tabela 7 – Resumo dos artigos com aprendizado de máquina com Deep Learning .	37
Tabela 8 – Bases de Dados.	40
Tabela 9 – Resumo da arquitetura proposta com um tamanho de janela de $25 \times$ 25. A última camada é baseada no conjunto de dados Indian Pines. . .	44
Tabela 10 – Resultados do Modelo Proposto nos Testes nas Diversas Bases de Dados	50
Tabela 11 – Comparação de desempenho entre os Modelos 2D-CNN, 3D-CNN e o Modelo Proposto com 10% de amostras no treinamento	66
Tabela 12 – Comparação de desempenho entre os Modelos 2D-CNN, 3D-CNN e o Modelo Proposto com 10% de amostras no treinamento	67
Tabela 13 – Resultados do Modelo Treinado em Indian Pines e Testado em Salinas	68
Tabela 14 – Resultados do Modelo Treinado em Indian Pines e Testado em Salinas com Ajuste Fino	69
Tabela 15 – Tabela de comparação para a base Indian Pines	69
Tabela 16 – Tabela de comparação para a base Salinas	70
Tabela 17 – Tabela de comparação para a base Universidade de Pavia	70
Tabela 18 – Tabela de comparação para a base KSC	70
Tabela 19 – Tabela de comparação para a base Botswana	71
Tabela 20 – Tabela de comparação para a base WHU-Hi-LongKou	71
Tabela 21 – Tabela de comparação para a base WHU-Hi-HanChuan	71
Tabela 22 – Tabela de comparação para a base WHU-Hi-HongHu	72

Lista de abreviaturas e siglas

AA	Acurácia Média
CNN	Rede Neural Convolucional
F1	F1-Score
GIC	Grupo de Inteligência Computacional
HSI	High-resolution hyperspectral image
KNN	K-Vizinhos Mais Próximos
KSC	Kennedy Space Center
MLP	Multilayer Perceptrons
OA	Acurácia Global
PCA	Análise de Componentes Principais
Prec	Precisão
Rec	Recall
RNN	Redes Neurais Recorrente
Spec	Especificidade
SVM	Support Vector Machines
UAV	Veículo aéreo não tripulado

Sumário

1	INTRODUÇÃO	1
1.1	Objetivo Geral	2
1.2	Objetivos Específicos	2
1.3	Publicações	3
1.4	Estrutura do Trabalho	3
2	REFERENCIAL TEÓRICO	5
2.1	Sensoriamento remoto hiperespectral	5
2.2	Técnicas clássicas de Aprendizado de Máquina	6
2.3	Deep Learning	7
2.4	Deep Learning em Sensoriamento Remoto Agrícola	8
2.4.1	Convolução 3D	10
2.4.2	Convolução 2D	10
2.4.3	CNNs 3D vs. CNNs 2D	11
2.5	Bases de Imagens	12
2.5.1	Base Indian Pines	12
2.5.2	Base Salinas	13
2.5.3	Base Universidade de Pavia	14
2.5.4	Base Centro Espacial Kennedy (KSC)	16
2.5.5	Botswana	18
2.5.6	WHU-Hi-LongKou	20
2.5.7	WHU-Hi-HanChuan	20
2.5.8	WHU-Hi-LongHu	21
2.6	Considerações Finais	22
3	ESTADO DA ARTE	23
3.1	Classificação de Imagens Hiperespectrais Aplicadas ao Mapeamento de Atividades Agrícolas: Abordagens com Aprendizado de Máquina	23
3.1.1	SVM	23
3.1.2	Aprendizado Não Supervisionado e Algoritmos Híbridos	24
3.1.3	Seleção de Bandas e Algoritmos de Ensemble	25
3.1.4	Campos Aleatórios Condicionais (CRF)	25
3.1.5	Abordagens Baseadas em Codificação Esparsa de Tensores para Classificação Hiperespectral	26
3.2	Classificação de imagens hiperespectrais aplicadas ao mapeamento de atividades agrícolas: abordagens com Deep Learning	28

3.2.1	Redes Convolucionais com Autoatenção	28
3.2.2	Otimização e Modelos Híbridos em Redes Neurais Convolucionais	30
3.2.3	Redes Convolucionais e Transformadores	32
3.3	Considerações Finais	38
4	MATERIAIS E MÉTODOS	40
4.1	Bases de Dados Utilizadas	40
4.2	Modelo Proposto	41
4.3	Divisão dos Dados e Treinamento	44
4.4	Métricas de Avaliação Utilizadas	45
4.5	Considerações Finais	49
5	RESULTADOS E DISCUSSÃO	50
5.1	Resultados do Método Proposto	50
5.1.1	Indian Pines	51
5.1.2	Salinas	53
5.1.3	Universidade de Pavia	54
5.1.4	Centro Espacial Kennedy (KSC)	56
5.1.5	Botswana	58
5.1.6	WHU-Hi-LongKou	60
5.1.7	WHU-Hi-HanChuan	61
5.1.8	WHU-Hi-HongHu	63
5.2	Experimentos com o Modelo	65
5.2.1	Experimento I - Desempenho com 10% de Amostras no Treinamento	65
5.2.2	Experimento II - Análise Cruzada entre Conjuntos de Dados	66
5.2.2.1	Avaliação do Modelo: Treinamento com Indian Pines e Teste com Salinas	67
5.2.3	Experimento III - Avaliação Cruzada entre base de dados	68
5.2.3.1	Congelamento de Camadas	68
5.2.3.2	Fine-Tuning	68
5.3	Comparação com a Literatura	69
5.4	Limitações	72
6	CONCLUSÕES E TRABALHOS FUTUROS	74
Conclusão		74
	REFERÊNCIAS	76

1 Introdução

O sensoriamento remoto tem sido fundamental no monitoramento e na gestão de recursos naturais desde meados do século XX. Inicialmente, o sensoriamento multiespectral, que captura imagens em poucas bandas espectrais amplas, permitia a identificação de características gerais da superfície terrestre (RICHARDS; RICHARDS et al., 2022). Com o avanço da tecnologia, surgiu o sensoriamento hiperespectral, que registra centenas de bandas espectrais estreitas e contínuas, proporcionando uma discriminação mais precisa de materiais e condições na superfície. Essa capacidade é particularmente valiosa na agricultura, onde o mapeamento preciso de culturas é essencial para o planejamento e o monitoramento ambiental (YOU; SUN, 2022; PATEL; PATEL; KATHIRIA, 2022).

A demanda por práticas agrícolas sustentáveis tem impulsionado a busca por métodos inovadores de monitoramento. As imagens hiperespectrais HSI (High-Resolution Hyperspectral Image) emergem como uma ferramenta poderosa, capturando informações detalhadas que permitem a classificação precisa de culturas e a análise das propriedades físico-químicas do solo (ZHONG et al., 2020; KARTHIKEYAN; CHAWLA; MISHRA, 2020). No entanto, o processamento dessas imagens enfrenta desafios significativos, como a necessidade de métodos eficientes para classificação e a dependência de grandes volumes de dados de treinamento (SUDHA; VAIDEGHY et al., 2022; FARMONOV et al., 2023).

Nesse contexto, técnicas de aprendizado de máquina, especialmente redes neurais convolucionais (CNN - Rede Neural Convolucional), têm-se destacado. Enquanto as CNNs 3D extraem informações espectrais e espaciais de forma integrada, sua complexidade pode ser um obstáculo. Por outro lado, as CNNs 2D, embora mais simples, podem não capturar plenamente o contexto espectral (ROY et al., 2019; GE et al., 2020). Para superar essas limitações, este trabalho propõe uma abordagem híbrida, combinando camadas convolucionais 3D e 2D para otimizar a extração de características e melhorar a precisão na classificação de culturas.

Para enfrentar desafios, como a falta de testes em bancos de dados públicos reconhecidos, a restrição de métricas utilizadas, a dependência de extensos volumes de dados de treinamento e a demanda por modelos que se ajustam a diversas situações e escalas, foi desenvolvido o modelo proposto. Este modelo foi avaliado em 08 (oito) conjuntos de dados hiperespectrais presentes na literatura, evidenciando um desempenho superior em relação às abordagens tradicionais. A sua utilização proporciona novas possibilidades para a categorização e o mapeamento hiperespectral em contextos agrícolas, disponibilizando uma ferramenta eficaz para o planejamento e a administração do setor. Além disso, esta dissertação investiga os aspectos teóricos e práticos do processamento de imagens

hiperespectrais, considerando como principais metodologias de aprendizagem de máquina e aprendizado profundo. Em adição, aborda os desafios que são enfrentados e fundamenta uma base robusta para investigações e progressos futuros no campo.

A elaboração e a validação deste modelo visam não apenas melhorar a eficiência na classificação de imagens hiperespectrais, mas também ampliar as oportunidades de utilização de tecnologias no setor agrícola. A relevância desta pesquisa transcende a elaboração de um novo modelo; abrange uma análise crítica das metodologias empregadas, além de uma revisão abrangente das técnicas utilizadas, contribuindo para o enriquecimento do entendimento e promovendo progressos na ciência das imagens hiperespectrais. Entre as principais contribuições, ressaltam-se: (i) a implementação de um pré-processamento de imagens hiperespectrais que melhora a representação dos dados e padroniza sua estrutura para a etapa de classificação; (ii) Criação de um modelo que classifica e mapeia imagens hiperespectrais; (iii) Usa uma arquitetura compacta de redes neurais convolucionais, que diminui o custo computacional; (iv) um modelo com a capacidade de generalização para diversas bases de dados hiperespectrais, conferindo ao modelo maior robustez para diversas aplicações; e (v) a diminuição da dependência em grandes volumes de dados de treinamento, o que amplia a abordagem em contextos com dados restritos. Tais progressos podem exercer uma influência específica na administração agrícola. Quando aprimorado, o modelo pode fomentar práticas mais sustentáveis e eficazes. A sua utilização pode melhorar a eficiência no emprego de recursos como água e fertilizantes, causando o impacto ambiental e elevando a produtividade das atividades, o que, conseqüentemente, colabora para a segurança alimentar no âmbito global ([THENKABAIL; LYON; HUETE, 2018](#); [GEBBERS; ADAMCHUK, 2010](#)).

1.1 Objetivo Geral

Este trabalho tem como objetivo propor um modelo de classificação e mapeamento de imagens hiperespectrais utilizando CNNs, por meio de uma combinação de arquiteturas 2D e 3D.

1.2 Objetivos Específicos

Para que o objetivo geral seja atingido, foram definidos os seguintes objetivos específicos:

- Revisar a literatura sobre técnicas de classificação de imagens hiperespectrais na agricultura, para identificar lacunas, oportunidades de inovação e limitações dos métodos atuais;

- Adquirir e explorar bases de imagens disponíveis de forma pública para o desenvolvimento da metodologia proposta;
- Desenvolver um método de pré-processamento que melhore a representação das imagens hiperespectrais, facilitando sua estruturação e otimizando a etapa de classificação;
- Projetar um modelo de CNN que integra camadas convolucionais 3D e 2D para a extração e processamento eficientes de características espectrais e espaciais;
- Validar o modelo proposto em relação a outras metodologias apresentadas na literatura, com variações em ajuste de camadas e distribuição de dados para treinamento e teste, a fim de obter maior capacidade de generalização.

1.3 Publicações

Esta dissertação resultou na elaboração de dois artigos submetidos ao *Journal of the Brazilian Computer Society* (JBCS). O primeiro artigo, intitulado "Image Classification of Hyperspectral Remote Sensing in Agricultural Activity Mapping: A Survey", destaca os principais resultados de uma revisão bibliográfica realizada durante o desenvolvimento desta pesquisa, refletindo tanto o impacto acadêmico quanto as aplicações práticas do modelo. O segundo artigo, intitulado "Hybrid Approach of 2D and 3D Convolutional Neural Networks for Classification and Mapping of Hyperspectral Images in Agriculture", apresenta os resultados obtidos neste trabalho, incluindo o desenvolvimento e a avaliação do modelo proposto.

1.4 Estrutura do Trabalho

O restante deste trabalho está organizado da seguinte forma:

- Capítulo 2 – Estado da Arte: Este capítulo apresenta um levantamento do estado da arte e das obras relacionadas ao tema proposto, destacando as principais contribuições e limitações dos trabalhos existentes.
- Capítulo 3 – Referencial Teórico: Aborda os conceitos fundamentais necessários para o entendimento do problema e das técnicas empregadas no desenvolvimento do trabalho, fornecendo a base teórica para as metodologias aplicadas.
- Capítulo 4 – Materiais e Métodos: Detalha a metodologia proposta neste trabalho, incluindo a descrição dos dados, as técnicas de processamento de imagens hiperespectrais utilizadas e os métodos de validação.

- Capítulo 5 – Resultados e Discussões: Apresenta os resultados alcançados com o método proposto e realiza uma comparação crítica com os resultados do estado da arte, seguida de uma discussão sobre os significados e implicações desses resultados.

- Capítulo 6 – Conclusões e Trabalhos Futuros: Sumariza as principais conclusões do trabalho, destacando as contribuições para o campo do sensoriamento remoto aplicado à agricultura e delineando sugestões para pesquisas futuras e possíveis melhorias no modelo desenvolvido.

2 Referencial Teórico

No capítulo atual, apresentamos uma fundamentação teórica sobre os fundamentos do processamento de imagens hiperespectrais e do aprendizado de máquina, com foco especial em técnicas de aprendizado profundo. Esta pesquisa emprega técnicas de CNN, em três dimensões (3D) e duas dimensões (2D), para enfrentar desafios complexos de classificação e mapeamento de áreas agrícolas por meio de imagens hiperespectrais. As redes neurais profundas são ideais para gerenciar a dimensionalidade e a vasta informação espectral e espacial presente nessas imagens, possibilitando a extração eficaz de atributos e uma classificação mais precisa. Este trabalho foca no modelo que integra camadas 3D e 2D numa estrutura para maximizar a eficiência na categorização e no mapeamento de zonas agrícolas. Este modelo busca explorar as capacidades do aprendizado profundo para aprimorar tanto a precisão quanto a utilidade prática das classificações em contextos de sensoriamento remoto agrícola, contribuindo significativamente para a agricultura de precisão e o manejo sustentável de recursos agrícolas.

2.1 Sensoriamento remoto hiperespectral

O sensoriamento remoto multiespectral teve início nos anos 70 com satélites como o Landsat, que coletavam informações em poucas faixas do espectro eletromagnético, possibilitando a avaliação de vastas regiões para monitoramento ambiental e agrícola. Contudo, a restrição na resolução espectral tornava a detecção de materiais mais específicos mais difícil. A progressão para o sensoriamento hiperespectral, que registra centenas de bandas espectrais finas, proporcionou maior exatidão, possibilitando a identificação minuciosa de diversos materiais e condições ambientais. Apesar de ser mais complexo e dispendioso, o sensoriamento hiperespectral se sobressai pela habilidade de realizar uma análise minuciosa e exata em relação ao multiespectral (RICHARDS; RICHARDS et al., 2022).

O sensoriamento remoto hiperespectral é fundamental para o mapeamento das atividades agrícolas, fornecendo dados espectrais minuciosos em várias bandas estreitas. Esta tecnologia é crucial para o reconhecimento e acompanhamento de vários parâmetros agrícolas, tais como tipos de plantações, estado de saúde e fases de desenvolvimento com exatidão. As imagens hiperespectrais, ao registrar assinaturas espectrais, auxiliam na distinção entre variados tipos de culturas e vegetação, possibilitando que os pesquisadores façam escolhas informadas sobre práticas de gestão de culturas, distribuição de recursos e monitoramento ambiental. (LU et al., 2020; THENKABAIL; LYON; HUETE, 2018; SU, 2020). Diferentemente de imagens multiespectrais, que capturam dados em poucas

bandas, as imagens hiperespectrais cobrem variações sutis na cor e morfologia das plantas, podendo identificar variações químicas, como níveis de nitrogênio e água, oferecendo uma ferramenta valiosa para práticas de agricultura de precisão. Além disso, a capacidade em produções de commodities (arroz, milho, algodão, dentre outros) de observar a reflectância em bandas específicas permite aos pesquisadores estudar fenômenos ambientais complexos, como estresse hídrico e infestação de pragas, com maior antecedência (NIGAM et al., 2019).

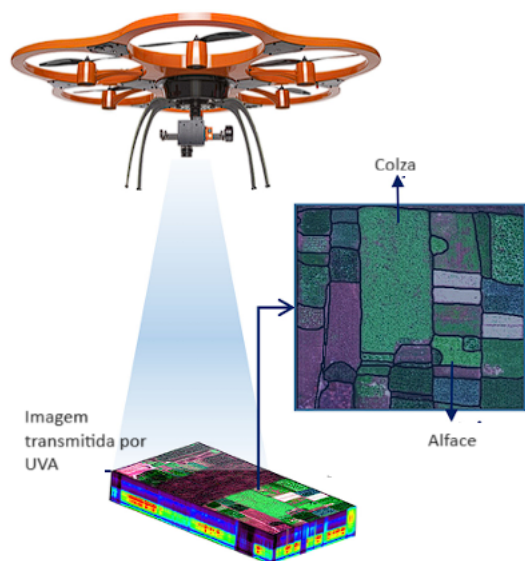


Figura 1 – Exemplo de captura de imagem Hiperespectral. Imagem adaptada do trabalho de Zhong et al. (2020).

A aquisição de imagens hiperespectrais, ilustrada na Figura 1, é frequentemente realizada por UAVs (Veículos Aéreos Não Tripulados) equipados com sensores avançados. Esta plataforma permite uma coleta de dados rápida e eficiente, embora apresente desafios como a limitação da capacidade de carga do UAV e a necessidade de condições meteorológicas ideais para a coleta eficaz de dados. Além disso, o processamento e a análise de grandes volumes de dados hiperespectrais requerem técnicas avançadas de processamento de imagens e aprendizado de máquina, destacando a importância de desenvolver algoritmos que possam eficientemente processar e interpretar essas informações complexas (YANG; HU, 2022).

A tecnologia hiperespectral também está evoluindo com o desenvolvimento de novos algoritmos de compressão e análise de dados que permitem uma gestão dos dados coletados, reduzindo a quantidade de dados transmitidos sem comprometer a qualidade das informações. Esses avanços são cruciais para otimizar as operações de coleta e análise, tornando o sensoriamento remoto hiperespectral ainda mais acessível e aplicável em contextos agrícolas globais (ZHONG et al., 2020).

2.2 Técnicas clássicas de Aprendizado de Máquina

As técnicas clássicas de Aprendizado de Máquina desempenham um papel fundamental no mapeamento de atividades agrícolas, oferecendo abordagens robustas e bem estabelecidas para analisar e interpretar dados agrícolas. Algoritmos de classificação, como Support Vector Machines (SVM), Árvores de Decisão, K-Vizinhos Mais Próximos (KNN) e Multilayer Perceptrons (MLP) (VAISHNAVI et al., 2022; WU et al., 2022), são comumente empregados para identificar padrões em imagens de sensoriamento remoto e mapear diferentes culturas e uso da terra (ZHANG et al., 2020).

Essas metodologias são apreciadas pela sua eficácia em gerenciar grandes volumes de dados e por proporcionar resultados de interpretação intuitiva, fatores fundamentais para a tomada de decisões na área. Adicionalmente, quando associadas a técnicas de pré-processamento de dados, como a Análise de Componentes Principais (PCA) e a transformação wavelet, essas abordagens têm o potencial de superar obstáculos ligados à dimensionalidade dos dados hiperespectrais, preservando dados cruciais para uma classificação precisa. Por exemplo, a Análise de Variância (PCA) é eficiente na diminuição da dimensionalidade dos dados, preservando os componentes mais significativos que capturam a maior variância. Por outro lado, a Transformação Wavelet pode decompor os dados em componentes de frequência, auxiliando na detecção de características delicadas que podem não ser perceptíveis em dimensões maiores (ZHANG et al., 2022; GALODHA et al., 2022a).

O uso combinado dessas técnicas clássicas com métodos de pré-processamento modernos aumenta a precisão da classificação, melhorando a eficiência computacional, tornando o processamento de grandes volumes de dados hiperespectrais mais viável e acessível para aplicações práticas em agricultura de precisão.

2.3 Deep Learning

O aprendizado profundo constitui uma vertente do aprendizado de máquina que utiliza redes neurais artificiais com múltiplas camadas para a modelagem de dados complexos e a realização de tarefas como classificação, regressão, detecção de objetos, tradução de idiomas, entre outras. Essa área consolidou-se como uma técnica eficiente devido à sua capacidade de aprender representações hierárquicas de dados, o que se revela especialmente vantajoso em aplicações como reconhecimento de imagens, processamento de linguagem natural e análise de dados hiperespectrais. A arquitetura do aprendizado profundo permite que os modelos identifiquem atributos em múltiplos níveis de abstração, variando desde elementos fundamentais nas camadas inferiores até padrões mais elaborados nas camadas superiores, resultando em um aprendizado mais profundo e robusto (SHINDE; SHAH, 2018).

Um dos elementos fundamentais das arquiteturas de aprendizado profundo são as Redes Neurais Multicamadas (MLPs) e as CNNs. As MLPs são constituídas por uma camada de entrada, diversas camadas ocultas e uma camada de saída. Todo neurônio em uma camada está integralmente interligado a todos os neurônios da camada subsequente, sendo que cada conexão apresenta um peso que pode ser ajustado. As camadas ocultas empregam funções de ativação não lineares, como a ReLU (Unidade Linear Retificada), sigmoid ou tanh, com o objetivo de incorporar não linearidade ao modelo, o que possibilita a captura de padrões complexos presentes nos dados. O treinamento destas redes consiste na minimização de uma função de perda, como a entropia cruzada, por meio da aplicação do algoritmo de retropropagação, o qual ajusta iterativamente os pesos e viés dos neurônios, com o objetivo de diminuir o erro nas previsões (WU et al., 2022).

A utilização de deep learning abrange múltiplas áreas, destacando-se a agricultura, na qual métodos de deep learning, particularmente os que empregam imagens hiperespectrais, demonstram potencial específico (PATEL; PATEL; KATHIRIA, 2022). Por meio de UAVs equipados com sensores hiperespectrais, é viável obter informações detalhadas sobre as atividades agrícolas, que podem ser analisadas por modelos de aprendizado profundo para a classificação de culturas, a identificação de patologias vegetais e o mapeamento de regiões agrícolas. Essas metodologias possibilitam uma análise mais precisa e eficaz, favorecendo a adoção de práticas agrícolas sustentáveis e produtivas. Além disso, análises comparativas entre MLPs e outros algoritmos, como KNN e SVM, evidenciaram que, apesar de MLPs e demais redes neurais profundas demandarem um maior volume de dados e capacidade de processamento, elas frequentemente superam outras abordagens em tarefas complexas, devido à sua competência para aprender representações ricas e hierárquicas dos dados (HU et al., 2022b).

2.4 Deep Learning em Sensoriamento Remoto Agrícola

As técnicas de Deep Learning, particularmente as CNNs, têm revolucionado o mapeamento de atividades agrícolas ao oferecer métodos poderosos para lidar com a complexidade e a variabilidade dos dados agrícolas. Estas redes são notáveis pela sua habilidade em processar e analisar eficazmente dados em formatos de imagem, permitindo a identificação precisa de diferentes culturas e características agrícolas. Arquiteturas como ResNet e DenseNet se destacam por suas profundas camadas de aprendizado, que são capazes de extrair características espaciais e espectrais complexas das imagens (GALODHA et al., 2022b; AGILANDEESWARI et al., 2022; BHOSLE; MUSANDE, 2022).

Tais redes são especialmente eficazes devido à sua capacidade de aprender automaticamente os filtros ótimos durante o processo de treinamento, em vez de depender de filtros manualmente projetados, o que é comum em métodos de aprendizado de máquina

tradicionais. A profundidade dessas arquiteturas permite que elas capturem uma hierarquia de características, das mais simples às mais complexas, o que é crucial para entender e classificar com precisão as diversas e frequentemente sutis diferenças entre tipos de culturas e condições de crescimento (AGILANDEESWARI et al., 2022; BHOSLE; MUSANDE, 2022).

A aplicação de sensoriamento remoto na agricultura possibilita uma administração mais eficaz e sustentável dos trabalhos. Com o auxílio de tecnologias como a inteligência artificial e o aprendizado de máquina, é possível analisar os dados obtidos para considerar padrões e antecipar problemas antes que estes se tornem críticos. Essa abordagem contribui para a elevação da produtividade, ao mesmo tempo em que diminui a utilização de recursos como água, fertilizantes e pesticidas, promovendo a adoção de práticas agrícolas mais sustentáveis. A evolução de abordagens como o aprendizado de transferência possibilitou a adaptação de modelos pré-treinados em extensos conjuntos de dados para especificações específicas de sensoriamento remoto, específicas a exigências de grandes volumes de dados agrícolas anotados. Essa metodologia diminui os custos e o tempo exigido para a elaboração de modelos eficientes, aprimorando a generalização desses modelos em diversos contextos de sensoriamento remoto, ou que amplia suas possíveis aplicações (GALODHA et al., 2022b).

As CNNs são essenciais no processamento e análise de imagens hiperespectrais devido à sua capacidade de aprender representações hierárquicas de características. Essa capacidade é útil para interpretar as complexas imagens hiperespectrais usadas em sensoriamento remoto agrícola. Além disso, a combinação entre RNNs, CNN Conv-4, CNN Conv-6, SSRN e CNNs 2D/3D está revolucionando a análise de séries temporais de imagens, permitindo detectar mudanças sazonais e prever rendimentos agrícolas, fornecendo informações sobre o desenvolvimento das culturas e as condições ambientais ao longo do tempo (YANG; HU, 2022; VAISHNAVI et al., 2022; PATEL et al., 2023; GALODHA et al., 2022a).

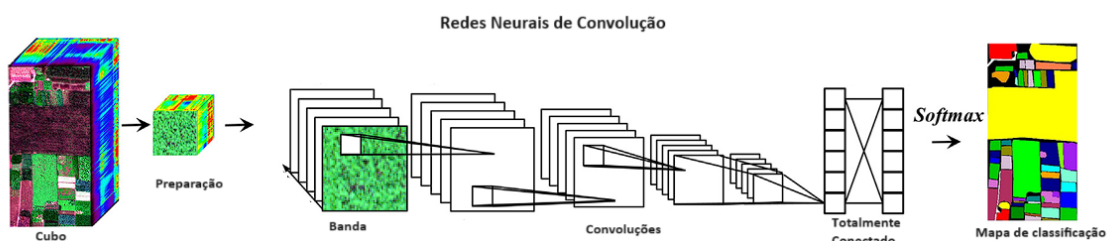


Figura 2 – CNN. Imagem adaptada do trabalho de ((ZHONG et al., 2020))

A Figura 2 ilustra a estrutura e o processo de uma CNN aplicada à classificação de imagens hiperespectrais, conforme adaptado do trabalho de Zhong et al. (2020). Inicialmente, observa-se um cubo de dados hiperespectrais, que é uma representação tridimensional onde cada pixel contém um espectro completo de informações em várias bandas de comprimento de onda, capturando detalhes ricos sobre as propriedades espectrais e

espaciais das áreas agrícolas. Esses dados são então preparados, passando por uma etapa de pré-processamento para extrair cubos menores representando regiões específicas da imagem original, facilitando o manuseio e processamento subsequente nas camadas convolucionais.

Os dados em forma de cubos são expostos a múltiplas camadas convolucionais, nas quais os filtros são usados para extrair características graves das imagens, tais como bordas, texturas e padrões. Depois das camadas convolucionais e de pooling, que diminuem a dimensionalidade dos dados enquanto preservam as características mais importantes, os dados são destinados para camadas totalmente conectadas. Nas referidas camadas, as particularidades extraídas são integradas para fazer a classificação. A camada final, denominada camada de Softmax, transforma os valores das saídas em probabilidades correspondentes a cada classe, gerando um mapa de classificação que associa uma classe específica a cada pixel da imagem original, como diferentes tipos de vegetação ou solo. Tal abordagem permite uma avaliação minuciosa e exata das zonas agrícolas, sendo eficaz para a consideração de variedades de cultivos, acompanhando o estado de saúde das vegetações e identificando pragas ou enfermidades, evidenciando a relevância das Redes Neurais Convolucionais (CNNs) para as inovações na agricultura de precisão.

2.4.1 Convolução 3D

A convolução tridimensional é aplicada sobre dados volumétricos, como cubos hiperespectrais, onde as dimensões são representadas pela altura, largura e profundidade (número de bandas espectrais ou frames temporais) (ZHONG et al., 2020). Para uma entrada $X \in \mathbb{R}^{H \times W \times D}$, onde H é a altura, W é a largura e D é a profundidade (bandas espectrais), a convolução 3D com um kernel $K \in \mathbb{R}^{k_H \times k_W \times k_D}$ pode ser representada pela seguinte equação:

$$Y(i, j, k) = \sum_{m=0}^{k_H-1} \sum_{n=0}^{k_W-1} \sum_{p=0}^{k_D-1} X(i+m, j+n, k+p) \cdot K(m, n, p) + b$$

Onde:

- $Y(i, j, k)$ é o valor da saída da convolução na posição (i, j, k) ,
- $X(i+m, j+n, k+p)$ é o valor da entrada na posição $(i+m, j+n, k+p)$,
- $K(m, n, p)$ é o valor do kernel na posição (m, n, p) ,
- b é o bias associado ao filtro.

Esta operação de convolução é aplicada de forma tridimensional, capturando tanto as informações espectrais quanto as espaciais dos dados de entrada (GOODFELLOW; BENGIO; COURVILLE, 2016).

2.4.2 Convolução 2D

A convolução bidimensional é aplicada sobre dados com duas dimensões, como imagens ou patches 2D extraídos de cubos hiperespectrais. Para uma entrada $X \in \mathbb{R}^{H \times W}$, onde H é a altura e W é a largura, a convolução 2D com um kernel $K \in \mathbb{R}^{k_H \times k_W}$ pode ser representada pela equação:

$$Y(i, j) = \sum_{m=0}^{k_H-1} \sum_{n=0}^{k_W-1} X(i+m, j+n) \cdot K(m, n) + b$$

Onde:

- $Y(i, j)$ é o valor da saída da convolução na posição (i, j) ,
- $X(i+m, j+n)$ é o valor do pixel da entrada na posição $(i+m, j+n)$,
- $K(m, n)$ é o valor do kernel na posição (m, n) ,
- b é o bias associado ao filtro.

Esta operação é aplicada sobre as duas dimensões espaciais, permitindo que o modelo capture as características visuais da entrada de forma eficiente (GOODFELLOW; BENGIO; COURVILLE, 2016).

2.4.3 CNNs 3D vs. CNNs 2D

O processamento de imagens hiperespectrais pode ser efetuado utilizando tanto redes convolucionais tridimensionais (3D) quanto bidimensionais (2D). As CNNs 3D são importantes por sua habilidade em capturar simultaneamente dependências espaciais e espectrais, processando dados em três dimensões (x, y e z, onde z representa o espectro). Esta capacidade é crucial em sensoriamento remoto, onde cada pixel contém uma rica informação espectral. Em contraste, CNNs 2D focam mais nas relações espaciais ao processar as imagens camada por camada, o que pode resultar em uma análise menos detalhada das complexidades integradas nas imagens hiperespectrais, embora sejam mais eficientes computacionalmente e mais simples de treinar (ROY et al., 2019).

Para maximizar os benefícios, a integração de CNNs 3D e 2D em uma arquitetura híbrida, como no modelo proposto, busca combinar a profunda interpretação das informações espectrais pelas CNNs 3D com a eficiência das CNNs 2D em processar características espaciais. Esta abordagem oferece uma solução robusta e eficiente para a classificação e o mapeamento hiperespectral de áreas agrícolas, permitindo uma análise detalhada e precisa dos vastos volumes de dados gerados por imagens hiperespectrais.

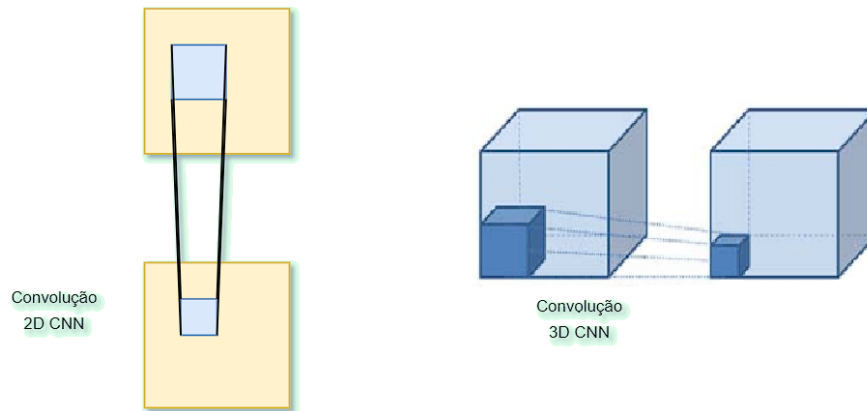


Figura 3 – Representação de convolução padrão 2D e 3D. Imagem adaptada do site ((<https://paperswithcode.com/method/3d-convolution>))

A Figura 3 demonstra claramente as diferenças entre as operações de convolução utilizadas nas CNNs 2D e 3D. No lado esquerdo da imagem, a representação mostra a convolução 2D, onde um filtro azul desliza sobre uma única dimensão espacial da entrada, que é representada em amarelo. Esta visualização enfatiza que a extração de características ocorre considerando as relações espaciais em duas dimensões, altura e largura. As setas pretas indicam como o filtro é aplicado horizontalmente ao longo da imagem, processando informações em duas dimensões. No lado direito, por outro lado, a convolução 3D é ilustrada com um filtro que se estende por três dimensões, aplicado sobre uma entrada também tridimensional. Esta abordagem é essencial onde a terceira dimensão pode ser o tempo ou a profundidade espectral (DIAKITE; JIANGSHENG; XIAPING, 2021).

2.5 Bases de Imagens

2.5.1 Base Indian Pines

A cena Indian Pines foi capturada pelo sensor AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) na região de Indian Pines, localizada no noroeste de Indiana, EUA. Ela abrange imagens que têm uma resolução de 145×145 pixels, com 200 bandas de refletância espectral, cobrindo comprimentos de onda entre 0,4 e $2,5 \mu\text{m}$. A área mapeada é composta principalmente por dois terços de terras agrícolas e um terço de floresta ou outra vegetação perene natural. Além disso, a cena inclui elementos antropogênicos, como duas rodovias de pista dupla, uma linha ferroviária, habitações de baixa densidade e outras infraestruturas menores. A imagem foi capturada no mês de junho, quando culturas como milho e soja estavam em estágios iniciais de crescimento, cobrindo menos de 5% da área total. O conjunto de dados contém 10.249 amostras, distribuídas em 16 classes distintas. A verdade terrestre associada a este conjunto de dados é categorizada nas seguintes classes (BASCO, 2021).

A Figura 4 representa o mapa de verdade terrestre (ground truth) do conjunto

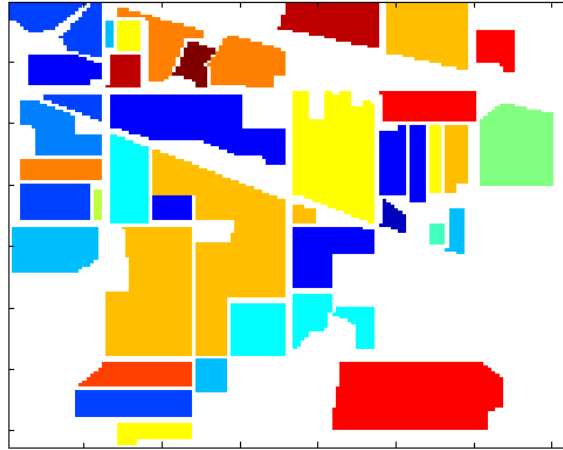


Figura 4 – Verdade terrestre do conjunto de dados Indian Pines

de dados Indian Pines, um dos mais utilizados na área de processamento de imagens hiperespectrais. Cada cor distinta indica uma das diferentes classes de cobertura do solo presentes na cena, como vegetação, floresta e áreas urbanizadas. Essa segmentação é essencial para avaliar a acurácia de modelos de classificação supervisionada, uma vez que fornece os rótulos de referência utilizados na validação dos algoritmos. A distribuição das classes não é homogênea, refletindo a variabilidade natural da região e impondo desafios adicionais para técnicas de aprendizado de máquina e redes neurais profundas, como a CNN 3D-2D empregada neste estudo.

Tabela 1 – Classes da verdade terrestre do conjunto de dados Indian Pines

Nº	Classes
1	Alfafa
2	Milho sem cultivo
3	Milho com cultivo mínimo
4	Milho
5	Pastagem
6	Gramma com árvores
7	Pastagem roçada
8	Feno enfardado
9	Aveia
10	Soja sem cultivo
11	Soja com cultivo mínimo
12	Soja limpa
13	Trigo
14	Floresta
15	Edificações com grama e árvores
16	Torres de aço e pedra

A Tabela 1 apresenta as 16 classes de cobertura do solo identificadas no conjunto de dados Indian Pines. Essas classes representam diferentes tipos de vegetação, áreas agrícolas e elementos artificiais dentro da cena capturada. Entre as categorias agrícolas,

destacam-se culturas como milho, soja e trigo, subdivididas conforme o tipo de manejo (exemplo: cultivo mínimo ou sem cultivo). Além disso, há classes relacionadas a vegetação natural, como pastagem e floresta, e também elementos não vegetais, como edificações e torres de aço. A presença de classes com características espectrais semelhantes torna o problema de classificação desafiador, exigindo abordagens avançadas, como redes neurais convolucionais 3D-2D, para uma segmentação precisa.

2.5.2 Base Salinas

A Figura 5 apresentada representa a verdade terrestre do conjunto de dados Salinas, um dos conjuntos de imagens hiperespectrais amplamente utilizados para classificação de uso do solo. Cada cor distinta na imagem indica uma classe diferente de cobertura do terreno, correspondendo a diferentes tipos de vegetação e superfícies encontradas na cena.

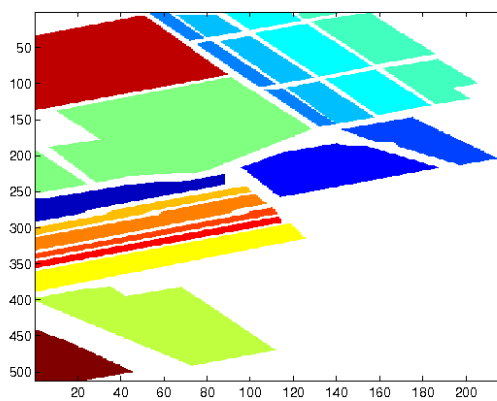


Figura 5 – Verdade terrestre do conjunto de dados Salinas

A cena Salinas foi coletada pelo sensor AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) de 204 bandas espectrais no Vale de Salinas, Califórnia, EUA. As imagens têm uma resolução de 512×217 pixels e estão disponíveis como dados de radiância no sensor. O conjunto de dados cobre uma área agrícola que inclui vegetais, solos nus e campos de vinhas. A verdade terrestre associada a este conjunto de dados é categorizada em 16 classes distintas, com um total de 54.129 amostras (BASCO, 2021).

A Tabela 2 apresenta as classes da verdade terrestre do conjunto de dados Salinas, amplamente utilizado na área de classificação de imagens hiperespectrais. Esse conjunto de dados contém 16 classes distintas, correspondentes a diferentes tipos de vegetação e superfícies agrícolas encontradas na região mapeada.

Observa-se que o conjunto de dados Salinas é particularmente relevante para aplicações em monitoramento agrícola, pois contém classes detalhadas relacionadas a culturas como alface romana, brócolis, milho e uvas, além de diferentes tipos de terreno em pousio e restolho. A presença de vinhedos com diferentes tipos de manejo também

torna esse conjunto interessante para estudos sobre o impacto das práticas agrícolas na variabilidade espectral.

A distinção entre classes como alface romana em diferentes estágios de crescimento (4 a 7 semanas) adiciona um desafio adicional para algoritmos de classificação supervisionada, pois demanda a identificação de variações sutis nos espectros de refletância ao longo do tempo. Além disso, a separação entre vinhedos treinados e não treinados pode ser útil para avaliações sobre a influência do manejo agrícola na resposta espectral das culturas.

O entendimento dessas classes é fundamental para o desenvolvimento de modelos de aprendizado de máquina e técnicas de processamento de imagens hiperespectrais, permitindo aplicações práticas em agricultura de precisão e monitoramento ambiental.

Tabela 2 – Classes da verdade terrestre do conjunto de dados Salinas

Nº	Classes
1	Brócolis (ervas daninhas 1)
2	Brócolis (ervas daninhas 2)
3	Terreno em Pousio
4	Terreno em Pousio (arado grosso)
5	Terreno em Pousio (suave)
6	Restolho
7	Aipo
8	Uvas
9	Solo de vinhedo em desenvolvimento
10	Milho (ervas daninhas)
11	Alface romana (4 semanas)
12	Alface romana (5 semanas)
13	Alface romana (6 semanas)
14	Alface romana (7 semanas)
15	Vinhedo (não treinado)
16	Vinhedo (treliça vertical)

2.5.3 Base Universidade de Pavia

Esta cena foi adquirida pelo sensor ROSIS durante uma campanha de voo sobre *Pavia*, no norte da Itália. O número de bandas é 103 para a *Pavia University* que tem imagens com resolução de 610×340 pixels. A verdade terrestre associada a esta imagem distingue 9 classes diferentes, com um total de 42.776 amostras (BASCO, 2021).

A Figura 6 apresenta a verdade terrestre (ground truth) do conjunto de dados hiperespectral da Universidade de Pavia. Essa verdade terrestre consiste em um mapa de referência, no qual diferentes classes de uso e cobertura do solo são representadas por cores distintas. Cada cor indica uma categoria específica, como áreas urbanas, vegetação e estradas, permitindo a validação de modelos de classificação supervisionados.

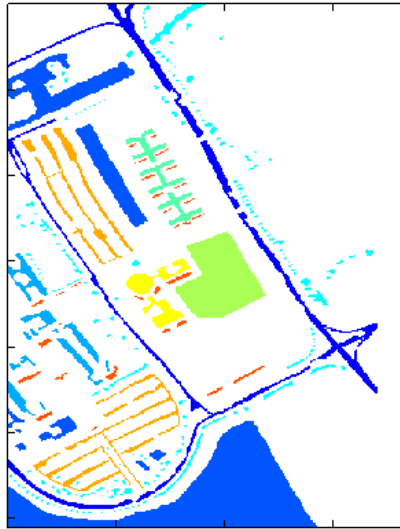


Figura 6 – Verdade terrestre do conjunto de dados Universidade de Pavia

Tabela 3 – Classes da verdade terrestre do conjunto de dados Universidade de Pavia

Nº	Classes
1	Asfalto
2	Prados
3	Cascalho
4	Árvores
5	Telhas metálicas pintadas
6	Solo exposto
7	Betume
8	Tijolos autoblocantes
9	Sombras

A Tabela 3 apresenta as classes de cobertura do solo presentes na verdade terrestre do conjunto de dados Universidade de Pavia. Esse conjunto de dados contém nove classes distintas, representando diferentes tipos de superfícies e materiais urbanos e naturais.

Cada classe corresponde a um tipo específico de cobertura do solo, incluindo superfícies pavimentadas, vegetação, materiais de construção e outros elementos como solo exposto e sombras. Essas categorias são essenciais para a validação de algoritmos de classificação de imagens hiperespectrais, permitindo avaliar a acurácia na distinção entre diferentes tipos de terreno e estruturas urbanas.

O conjunto de dados da Universidade de Pavia é amplamente utilizado na literatura devido à sua diversidade de classes e complexidade espectral, tornando-se uma referência para experimentos com técnicas de aprendizado profundo e aprendizado de máquina aplicadas à classificação de imagens hiperespectrais.

2.5.4 Base Centro Espacial Kennedy (KSC)

O instrumento NASA AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) adquiriu os dados no Kennedy Space Center (KSC), Flórida (EUA), em 23 de março de 1996. As imagens têm uma resolução de 512×614 pixels e contêm um total de 5211 amostras. A base de dados apresenta 176 bandas espectrais, cada uma com 10 nm de largura, cobrindo comprimentos de onda centrais que variam de 400 a 2500 nm. Os dados do KSC, adquiridos a uma altitude de aproximadamente 20 km, têm uma resolução espacial de 18 metros. Para fins de classificação, foram definidas 13 classes conforme Tabela 4 representativas dos diversos tipos de cobertura do solo que ocorrem neste ambiente (BASCO, 2021).

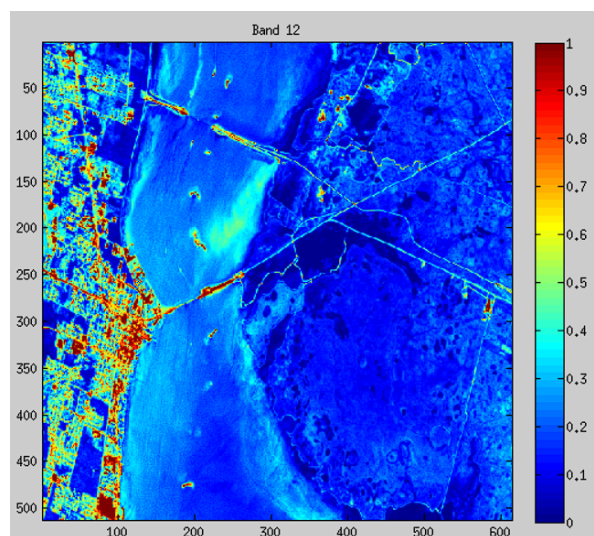


Figura 7 – Band 12 - KSC

A Figura 7 apresenta a imagem correspondente à Banda 12 do conjunto de dados hiperespectral Kennedy Space Center (KSC). Esse conjunto de dados contém uma diversidade de tipos de vegetação e superfícies naturais. Nesta imagem, as intensidades espectrais estão representadas por uma escala de cores, onde as regiões em azul indicam valores de reflectância mais baixos, enquanto as áreas em vermelho e amarelo representam valores mais altos. Essa variação na intensidade espectral reflete diferentes propriedades do solo e da cobertura do terreno, sendo essencial para a distinção entre classes em aplicações de classificação supervisionada e análise ambiental.

A Tabela 4 apresenta as 13 classes da base de dados hiperespectral Kennedy Space Center (KSC). Esse conjunto de dados é caracterizado por uma grande diversidade de tipos de vegetação e superfícies naturais. As classes dessa base de dados representam diferentes tipos de cobertura do solo, predominantemente formadas por vegetação, incluindo arbustos, diversos tipos de pântanos e formações arbóreas como pinheiros e carvalhos. Além disso, a base também inclui categorias como planícies de lama e corpos d'água, que adicionam complexidade ao problema de classificação de imagens hiperespectrais.

Tabela 4 – Classes da base de dados KSC

Nº	Classe
1	Arbustos
2	Pântano de Salgueiro
3	Mata de CP
4	Pinheiro Slash
5	Carvalho/Folhosa
6	Madeira Dura
7	Pântano
8	Pântano Graminóide
9	Pântano de Spartina
10	Pântano de Taboa
11	Pântano Salgado
12	Planícies de Lama
13	Água

A base de dados KSC é amplamente utilizada em pesquisas de aprendizado de máquina e aprendizado profundo devido à sua variedade de classes e à complexidade espectral associada à vegetação costeira. Esses fatores tornam esse conjunto de dados um desafio interessante para algoritmos de classificação e segmentação de imagens hiperespectrais.

2.5.5 Botswana

O satélite EO-1 da NASA adquiriu uma sequência de dados sobre o Delta do Okavango, em Botswana, durante os anos de 2001 a 2004. As imagens têm uma resolução de 1476×256 pixels e contêm um total de 3.248 amostras. A base final apresenta 145 bandas espectrais, cobrindo a faixa de 400 a 2500 nm do espectro em janelas de 10 nm. O pré-processamento dos dados foi realizado pelo Centro de Pesquisa Espacial da UT, com o objetivo de mitigar os efeitos de detectores defeituosos, calibração incorreta entre detectores e anomalias intermitentes. Foram definidas 14 classes que representam os tipos de cobertura da terra em pântanos sazonais, pântanos ocasionais e florestas mais secas localizadas na porção distal do Delta ([BASCO, 2021](#)).

A Figura 8 apresenta a imagem correspondente à Banda 101 do conjunto de dados hiperespectral Botswana. A imagem exibe a variação espectral da cena, com uma escala de cores indicando diferentes intensidades de reflectância. Regiões em azul representam áreas com menor resposta espectral nessa banda, enquanto tons de amarelo e vermelho indicam maior reflectância. Essas variações estão associadas a diferentes tipos de vegetação, corpos d'água e solos presentes na região.

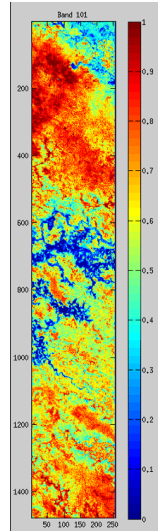


Figura 8 – Band 101 - Botswana

A Tabela 5 apresenta as 14 classes da base de dados hiperespectral Botswana. Esse ecossistema é caracterizado por uma grande diversidade de tipos de vegetação e superfícies naturais, tornando esse conjunto de dados uma referência importante para estudos ambientais e análise de imagens hiperespectrais.

Tabela 5 – Classes da base de dados Botswana

Nº	Classe
1	Água
2	Gramma de hipopótamo
3	Gramíneas de planície aluvial 1
4	Gramíneas de planície aluvial 2
5	Juncos
6	Vegetação ripária
7	Cicatriz de fogo
8	Interior da ilha
9	Florestas de acácia
10	Terras arbustivas de acácia
11	Gramíneas de acácia
12	Mopane baixo
13	Mopane misto
14	Solos expostos

As classes presentes nesse conjunto de dados incluem diferentes formações vegetais, como gramíneas, juncos, florestas de acácia e vegetação ripária, além de categorias associadas a características ambientais específicas, como cicatrizes de fogo e solos expostos. A presença de corpos d'água também é registrada, sendo um fator relevante para o monitoramento da dinâmica hidrológica da região.

A base de dados Botswana é amplamente utilizada para o desenvolvimento e teste de algoritmos de classificação hiperespectral, aprendizado profundo e análise ambiental.

Sua complexidade espectral e a variabilidade nas respostas dos diferentes tipos de cobertura do solo tornam essa base um desafio significativo para a modelagem de padrões espectrais e a segmentação de classes em imagens hiperespectrais.

2.5.6 WHU-Hi-LongKou

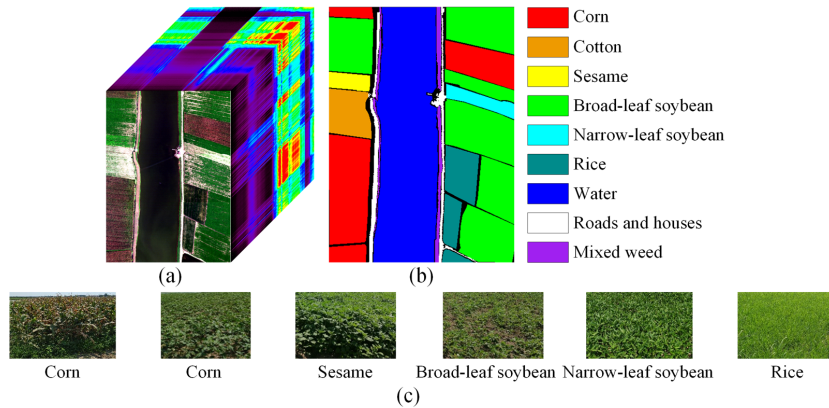


Figura 9 – (a) Cubo de imagem. (b) Verdade terrestre. (c) Fotos típicas das culturas na área de estudo. Fonte: (ZHONG et al., 2020)

Na Figura 9 temos um exemplo dos dados do conjunto WHU-Hi-LongKou que foram capturados entre as 13:49 e as 14:37 do dia 17 de julho de 2018, na cidade de Longkou, província de Hubei, China. Um sensor de imagem hiperespectral *Nano-Hyperspec* de 8 mm de distância focal estava instalado em uma plataforma UAV *DJI Matrice 600 Pro* (DJI M600 Pro). Durante a coleta, as condições climáticas eram favoráveis, com céu claro e temperatura em torno de 36°C e umidade relativa do ar de aproximadamente 65%. A área estudada é uma paisagem agrícola simples, abrangendo seis variedades de culturas: milho, algodão, gergelim, soja de folha larga, soja de folha estreita e arroz. O UAV voou a uma altitude de 500 m, com as imagens resultantes possuindo dimensões de 550×400 pixels e cobrindo 270 bandas espectrais de 400 a 1000 nm. A resolução espacial das imagens hiperespectrais capturadas pelo UAV é de cerca de 0,463 m. O conjunto de dados possui um total de 204.542 amostras distribuídas em 9 classes (ZHONG et al., 2020).

2.5.7 WHU-Hi-HanChuan

Na Figura 10 apresenta o conjunto de dados WHU-Hi-HanChuan que foi obtido das 17:57 às 18:46 do dia 17 de junho de 2016, na cidade de Hanchuan, província de Hubei, China. Utilizou-se um sensor de imagem hiperespectral *Nano-Hyperspec* de 17 mm de distância focal, montado em uma plataforma UAV *Leica Aibot X6 V1*. Durante a coleta, o clima estava claro e sem nuvens, com temperatura em torno de 30°C e umidade relativa do ar de aproximadamente 70%. A área estudada é uma zona de transição entre áreas

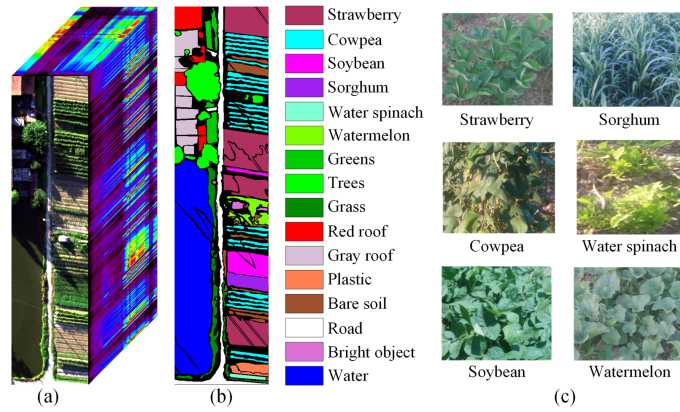


Figura 10 – (a) Cubo de imagem. (b) Verdade terrestre. (c) Fotos típicas das culturas na área de estudo. Fonte: (ZHONG et al., 2020)

rurais e urbanas, com edifícios, corpos d’água e terras cultivadas, abrangendo sete espécies de culturas: morango, feijão-caupi, soja, sorgo, espinafre d’água, melancia e verduras. O UAV operou a uma altitude de 250 m, gerando imagens de 1217×303 pixels, cobrindo 274 bandas espectrais de 400 a 1000 nm. A resolução espacial das imagens hiperspectrais capturadas pelo UAV é de cerca de 0,0109 m. Notavelmente, devido à aquisição durante a tarde, quando o ângulo de elevação solar era baixo, muitas áreas na imagem estão cobertas por sombras. Uma visão geral desses dados é mostrada na Figura 10. O conjunto de dados contém um total de 257.530 amostras distribuídas em 16 classes (ZHONG et al., 2020).

2.5.8 WHU-Hi-LongHu

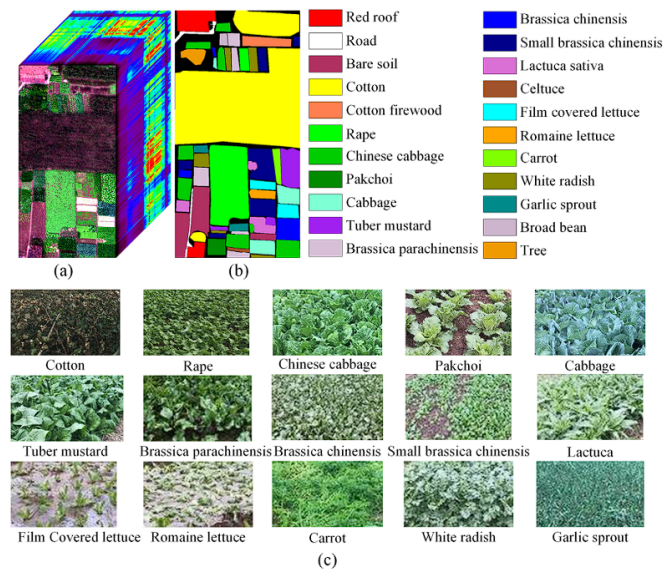


Figura 11 – (a) Cubo de imagem. (b) Verdade terrestre. (c) Fotos típicas das culturas na área de estudo. Fonte: (ZHONG et al., 2020)

Na Figura 11 temos o conjunto de dados WHU-Hi-HongHu que foi adquirido das 16h23 às 17h37 em 20 de novembro de 2017, na cidade de Honghu, província de Hubei, China, com uma distância focal de 17 mm. O *Headwall Nano-Hyperspec* sensor de imagem equipado em uma plataforma *DJI Matrice 600 Pro* UAV foi utilizado durante a coleta de dados. O tempo estava nublado, a temperatura estava em torno de 8°C e a umidade relativa do ar estava em torno de 55%. Classes de culturas e diferentes cultivares da mesma cultura também são plantadas na região, incluindo couve-chinesa e repolho, e *Brassica chinensis* e pequena *Brassica chinensis*. Notavelmente, a região é plantada com diferentes cultivares do mesmo tipo de cultura, por exemplo: couve-chinesa/repolho e *brassica chinensis*/pequena *brassica chinensis*. O UAV voou a uma altitude de 100 m, o tamanho da imagem é 940×475 pixels, existem 270 bandas de 400 a 1000 nm e a resolução espacial do UAV. As imagens hiperespectrais transmitidas são de cerca de 0,043 m. O conjunto de dados contém um total de 386.693 amostras distribuídas em 22 classes. (ZHONG et al., 2020).

2.6 Considerações Finais

Este capítulo fornece uma revisão do processamento e análise de imagens hiperespectrais, com um enfoque em suas aplicações na agricultura. A discussão transitou desde as técnicas clássicas de aprendizado de máquina até as inovações mais recentes em deep learning, destacando especialmente o papel das CNNs. Além disso, o próximo capítulo será sobre os materiais e métodos utilizados no desenvolvimento do modelo proposto. Este estudo destaca a relevância das inovações em deep learning para o sensoriamento remoto e enfatiza como essas tecnologias podem ser estrategicamente aplicadas para melhorar a produtividade e a eficiência na agricultura de precisão.

3 Estado da Arte

Tendo em vista as dificuldades na categorização de imagens hiperespectrais em práticas agrícolas, este capítulo oferece um panorama da atualidade sobre a temática. Uma metodologia de revisão bibliográfica além das fases delineadas por Kitchenham ([KITCHENHAM, 2004](#)), que compreende: planejamento, execução e seleção das informações. Nesta seção, são examinados os principais estudos que utilizam tanto abordagens tradicionais de aprendizagem de máquina quanto métodos de aprendizagem profunda para a análise e categorização dessas imagens. Ao investigar essas abordagens, objetivamos entender suas contribuições e limitações, bem como os progressos que elas viabilizaram, estabelecendo uma fundamentação robusta para a apresentação do modelo sugerido.

3.1 Classificação de Imagens Hiperespectrais Aplicadas ao Mapeamento de Atividades Agrícolas: Abordagens com Aprendizado de Máquina

Nesta seção, exploramos como técnicas clássicas de aprendizado de máquina têm sido aplicadas à classificação de imagens hiperespectrais no contexto agrícola. Abordamos desde métodos estabelecidos como SVM e Random Forest, que são otimizados para lidar com dados complexos, até algoritmos de aprendizado não supervisionado e híbrido, que oferecem soluções eficazes em cenários com poucos dados rotulados. Discutimos a importância da seleção de bandas e o uso de algoritmos de ensemble para melhorar a precisão, além da aplicação de Campos Aleatórios Condicionais (CRF - Conditional Random Fields) para incorporar dependências espaciais e espectrais. Finalmente, destacamos as abordagens novas baseadas em codificação esparsa de tensores, que preservam características essenciais das imagens e oferecem eficiência computacional significativa.

3.1.1 SVM

Baabu e Padma ([BAABU; PADMA, 2021](#)) introduziram o método SFR-HSI (*Semantic Feature Representation based HSI*), que utiliza SVM para a classificação de culturas hiperespectrais. Este método aborda desafios como a alta dimensionalidade, correlação espectral e presença de pixels mistos em imagens hiperespectrais, alcançando uma acurácia de 97,98% no banco de dados Indian Pines, superando outros modelos como LJCRC e MLJCRC. No entanto, o estudo destaca algumas limitações, como a sensibilidade de ruído nas imagens e a variabilidade das condições ambientais e de iluminação. Melhorias no método de redução dimensional para obter uma extração de características mais eficiente

e também o desenvolvimento do algoritmo de classificação aprimorado que aprenda pixels mistos de forma mais eficiente.

Zhang et al. (ZHANG et al., 2020) exploraram a importância do estágio de crescimento das plantas utilizando o modelo GA-SVM para a classificação de espécies vegetais, baseado em uma biblioteca espectral de plantas. O estudo alcançou uma acurácia de 97% em cenários agrícolas, demonstrando a eficácia da seleção e análise de robustez de características espectrais sensíveis. No entanto, a classificação de espécies vegetais baseada em imagens hiperespectrais reais de UAV também comprovou a eficácia deste método, embora a influência ambiental e o estado de crescimento das culturas tenham complicado a classificação, resultando em uma precisão significativamente melhor quando baseada em um lote inteiro em comparação com a classificação pixel a pixel. O estudo também destaca a necessidade de otimização dos algoritmos para melhorar a robustez e a aplicabilidade dos resultados em diferentes condições ambientais e estágios de crescimento das plantas.

3.1.2 Aprendizado Não Supervisionado e Algoritmos Híbridos

Nigam et al. (NIGAM et al., 2019) exploram o uso de imagens hiperespectrais para a discriminação de tipos de culturas e a avaliação da saúde vegetal. Utilizando o espectrômetro AVIRIS-NG, com cerca de 425 bandas espectrais, o estudo capturou imagens em várias regiões agrícolas da Índia. Os métodos de classificação empregados, como SAM (Spectral Angle Mapper) e MLC (Maximum Likelihood Classification), alcançaram precisões entre 77% e 94%. Foram realizadas análises de conteúdo de clorofila e índice de área foliar usando modelos de transferência radiativa de dossel, além de se explorar a discriminação do conteúdo de nitrogênio nas folhas e a identificação de estresses bióticos em culturas de trigo. Técnicas de redução de dimensionalidade como PCA (Análise de Componentes Principais) e transformação MNF (Maximum Noise Fraction) foram utilizadas para melhorar a precisão da classificação. Contudo, é válido destacar que não foram realizados testes em bases de dados públicas para avaliar o modelo e a necessidade também de avaliação entre bases cruzadas para avaliar melhor o teste.

Wu et al. (WU et al., 2023) exploraram o aprendizado não supervisionado para identificar infestações de ácaros-aranha em árvores de jujube, utilizando o algoritmo BC-DP. O método atingiu uma precisão de 96,13%, destacando a eficácia da fusão de informações espectrais e espaciais para melhorar a acurácia do reconhecimento. No entanto, o estudo reconhece várias limitações, como a necessidade de desenvolver sistemas de monitoramento de pragas tridimensionais, precisos e em tempo real, bem como a otimização de algoritmos de clustering para melhorar a eficiência computacional e a resistência ao ruído. Além disso, o uso de UAVs, embora altamente preciso, é limitado por fatores regionais e de custo.

Nandibewoor e Hegadi (NANDIBEWOOR; HEGADI, 2019) propuseram uma abordagem híbrida para segmentação de imagens hiperespectrais usadas para discriminar

tipos de culturas e avaliar a saúde vegetal em regiões agrícolas, com foco na estimativa do conteúdo de clorofila nas plantas, combinando Regressão Logística Multinomial Esparsa (SMLR) com Campos Aleatórios de Markov, otimizados por Enxame de Partículas (PSO), para estimar o conteúdo de clorofila nas culturas. O método demonstrou uma precisão de 93%, superando abordagens convencionais em termos de acurácia e tempo de execução. No entanto, ainda enfrenta desafios relacionados à necessidade de mais validação para garantir a generalização em diferentes condições de campo, além de requerer melhorias na otimização de parâmetros e no pré-processamento das imagens hiperespectrais para aumentar sua robustez e aplicabilidade em ambientes variados.

3.1.3 Seleção de Bandas e Algoritmos de Ensemble

Dave et al. (DAVE; VYAS; TRIVEDI, 2022) apresentaram o algoritmo BSSID (Band Selection using Spectral Information Divergence) para seleção de bandas em dados hiperespectrais, obtendo uma acurácia de 98,43% com KNN. O principal benefício do método proposto é sua rapidez computacional, permitindo a seleção de bandas usando apenas dois espectros. No entanto, o método apresenta limitações quanto à especificidade para os conjuntos de dados testados e não há melhorias substanciais além da seleção de 20 bandas, indicando que o desempenho adicional é mínimo com mais bandas.

No estudo de Colkesen e Ozturk (COLKESEN; OZTURK, 2022), avaliam-se algoritmos de aprendizado de ensemble para classificação de cobertura terrestre usando imagens de satélite de alta e média resolução, destacando-se os modelos CCF (Canonical Correlation Forest), LightGBM (Light Gradient Boosting Machine) e RotFor (Rotation Forest) (LIANG et al., 2020; DAOUD, 2019). O RotFor foi superior em imagens de alta resolução com acurácia global de 95,70%, seguido pelo CCF com 95,53%, mostrando eficácia na identificação de classes terrestres em imagens hiperespectrais. Apesar da eficácia desses modelos avançados, estes também enfrentam limitações como a especificidade para o conjunto de dados testado, usando um banco de dados público hiperespectral quando comparado a outros trabalhos. É válido destacar também a necessidade de testes com mais métricas de avaliação para avaliar a robustez dos modelos.

3.1.4 Campos Aleatórios Condicionais (CRF)

No artigo de Zhao et al. (ZHAO et al., 2020), destaca-se a importância do sensoriamento remoto de alta resolução via UAVs para o monitoramento de culturas em pequenas propriedades. O estudo apresenta o método SCRF (Spectral-Spatial Conditional Random Fields), que aplica Campos Aleatórios Condicionais para integrar informações espectrais e espaciais, melhorando a classificação de culturas (LAFFERTY et al., 2001). Experimentos indicam que o SCRF é superior a métodos tradicionais, corrigindo ruído em bandas contaminadas e aumentando a precisão da classificação. Em teste na base de

dados Salinas, obteve uma acurácia global de 99,40%. No entanto, o método depende de alta resolução das imagens, exigindo também extensivos dados de treinamento, o que pode ser desafiador em áreas com diversidade de culturas e dados limitados. Essas restrições apontam para a necessidade de futuras melhorias no SCRF, especialmente na adaptação para diferentes tipos de culturas, redução dos requisitos computacionais e desenvolvimento de abordagens mais eficazes para a seleção de amostras.

No estudo de Wei et al. (WEI et al., 2019b), é proposto o método SSF-CRF (Spatial-Spectral Fusion based on Conditional Random Fields), que combina características espectrais e espaciais para lidar com mudanças espectrais e heterogeneidade em áreas agrícolas complexas. Esse método melhora a precisão da classificação em imagens hiperespectrais de alta resolução, obtendo uma acurácia de 97,95% na base HongHu. O SSF-CRF se destaca por sua capacidade de integrar informações essenciais das imagens, mas ainda enfrenta desafios relacionados à generalização para diferentes tipos de culturas e condições agrícolas.

Em outro trabalho de Wei et al. (WEI et al., 2019a), é introduzido o método SSLF-CRF (Spectral-Spatial-Location Fusion based on Conditional Random Fields), que visa superar as limitações de abordagens tradicionais ao integrar informações espectrais, contextuais e espaciais. Essa abordagem proporciona uma melhoria significativa na precisão da classificação em larga escala, atingindo uma acurácia global de 98,07% na base de dados HongHu. Apesar dos resultados promissores, é necessário avaliar esses métodos em bancos de dados públicos adicionais para uma análise mais abrangente do desempenho, especialmente em termos de generalização para diferentes condições agrícolas.

3.1.5 Abordagens Baseadas em Codificação Esparsa de Tensores para Classificação Hiperespectral

No estudo de Yang et al. (YANG et al., 2023a), os autores propõem o classificador KTSSCC (Kernel Tensor Slice Sparse Coding-Based Classifier) para a classificação precisa de culturas em imagens hiperespectrais de UAVs, visando melhorar a discriminabilidade das imagens e preservando características espectrais e espaciais. Testado em dois conjuntos de dados reais, o KTSSCC superou métodos como SVM, CNN com CRF, SPNet e HPSCC-SNT. Utilizando uma representação de tensor espacial de vizinhança kernel (KSNT - Kernel Spatial Neighbor Tensor) e um algoritmo de busca ortogonal esparsa (KTSSOMP - Kernel Tensor Slice Sparse Orthogonal Matching Pursuit), o método reduz o custo computacional e aumenta a robustez. O KTSSCC apresentou desempenho superior quando comparado a outros modelos em termos de precisão geral (OA) com 98,07%. No entanto, o estudo reconhece limitações, como a necessidade de mais pesquisas sobre aprendizado por transferência para diferentes estágios de crescimento das culturas e a existência de necessidade de uma avaliação mais abrangente do desempenho do modelo em diferentes

cenários e condições.

No estudo de Yang et al. (YANG et al., 2023b), os autores propõem um modelo de codificação esparsa de tensor kernel (KTSCM - Kernel Tensor Sparse Coding Model) para a classificação precisa de culturas utilizando imagens hiperespectrais de UAVs. O método melhora a separação linear enquanto preserva as estruturas espectrais e espaciais das coberturas terrestres, reduzindo a complexidade computacional através de operações de matriz. Avaliado em dois conjuntos de dados hiperespectrais reais, o KTSCM superou métodos tradicionais e avançados, como SRC, WSGR, GTR, TLRSD-IDL e HPSCC-SNT, com precisão geral e média acima de 98,5% e valores kappa superiores a 0,98. Contudo, o estudo reconhece limitações, como a necessidade de mais pesquisas sobre o uso de aprendizado por transferência para diferentes estágios de crescimento das culturas. E também a etapa de pré-processamento não recebe a ênfase necessária, que poderia ser utilizada para melhorar o desempenho do modelo.

Tabela 6 – Resumo dos artigos sobre Aprendizado de Máquina

Artigo	Aplicação Investigada	Técnica Utilizada	Resultados	Bases de Dados
Baabu e Padma (BAABU; PADMA, 2021)	Classificação de culturas hiperespectrais	SFR-HSI com SVM	97,98% de acurácia	Indian Pines
Zhang et al. (ZHANG et al., 2020)	Classificação de espécies vegetais	GA-SVM	97% de acurácia	Hangzhou, China
Nigam et al. (NIGAM et al., 2019)	Discriminação de tipos de culturas e avaliação da saúde vegetal	SAM, MLC, PCA, MNF (AVIRIS-NG)	Precisões entre 77% e 94%	Regiões agrícolas homogêneas e heterogêneas na Índia
Wu et al. (WU et al., 2023)	Identificação de infestações de ácaros-aranha em jujube	BC-DP (aprendizado não supervisionado)	96,13% de precisão	Região de Hotan, Xinjiang, China
Nandibewoor e Hegadi (NANDIBEWOOR; HEGADI, 2019)	Segmentação de imagens hiperespectrais	SMLR + Campos Aleatórios de Markov	93% de precisão	Região de North Karnataka, Índia
Dave et al. (DAVE; VYAS; TRIVEDI, 2022)	Seleção de bandas em classificação de culturas	BSSID	98,43% de acurácia com KNN	Indian Pines e Salinas
Colkesen e Ozturk (COLKESEN; OZTURK, 2022)	Classificação de culturas hiperespectrais	CCF, LightGBM, RotFor	95,70% de acurácia com RotFor	Universidade de Pavia
Zhao et al. (ZHAO et al., 2020)	Classificação de culturas em pequenas propriedades	SCRF (Campos Aleatórios Condicionais)	99,40% de acurácia	Salinas e WHU-HiHanChuan
Wei et al. (WEI et al., 2019b)	Fusão espectral-espacial para classificação agrícola	SSF-CRF	97,95% de acurácia	Hanchuan e Honghu na província de Hubei, China
Wei et al. (WEI et al., 2019a)	Fusão espectral-espacial-local para classificação agrícola	SSLF-CRF	98,07% de acurácia	Hanchuan e Honghu na província de Hubei, China
Yang et al. (YANG et al., 2023a)	Classificação de culturas em imagens hiperespectrais de UAVs	KTSSCC (Kernel Tensor Slice Sparse Coding-Based Classifier)	98,07% de acurácia	Hanchuan e Honghu na província de Hubei, China
Yang et al. (YANG et al., 2023b)	Classificação de culturas em imagens hiperespectrais de UAVs	KTSCM (Kernel Tensor Sparse Coding Model)	98,5% de acurácia média, Kappa > 0.98	Hanchuan e Honghu na província de Hubei, China

A Tabela 6 apresenta um resumo dos principais trabalhos relacionados à aplicação de técnicas de aprendizado de máquina clássico e abordagens recentes para a classificação de

imagens hiperespectrais em contextos agrícolas. Os trabalhos listados na tabela demonstram a diversidade de técnicas aplicadas à classificação hiperespectral, desde métodos clássicos, como Support Vector Machines (SVM) e análise de componentes principais (PCA), até abordagens mais recentes, como Campos Aleatórios Condicionais (CRF) e modelos baseados em aprendizado profundo. Os resultados obtidos variam entre 77% e 99,40% de acurácia, dependendo da técnica e da complexidade do problema abordado. Destacam-se estudos como o de Zhao et al. (ZHAO et al., 2020), que alcançou 99,40% de acurácia utilizando SCRF (Spectral Conditional Random Fields), e o de Yang et al. (YANG et al., 2023b), que obteve 98,5% de acurácia média com o modelo KTSCM (Kernel Tensor Sparse Coding Model).

Além disso, a tabela evidencia a importância da escolha adequada das bases de dados, que variam desde conjuntos públicos amplamente utilizados, como Salinas, até dados específicos coletados por veículos aéreos não tripulados (UAVs). Essa variedade de bases de dados permite avaliar a robustez e a generalização dos modelos propostos em diferentes cenários agrícolas.

Em resumo, a Tabela 6 oferece uma visão abrangente do estado da arte em classificação hiperespectral aplicada à agricultura, destacando as técnicas mais promissoras e seus respectivos desempenhos. Esses resultados servem como base para a proposta deste trabalho, que busca avançar no desenvolvimento de modelos híbridos capazes de integrar características espectrais e espaciais para melhorar a precisão e a eficiência na classificação de culturas agrícolas. Assim, os artigos revisados não apenas ajudaram na definição do modelo proposto, mas também ofereceram uma compreensão abrangente dos avanços no campo, permitindo que o modelo fosse projetado de forma a capturar de maneira eficiente as múltiplas dimensões espectrais e espaciais dos dados agrícolas.

3.2 Classificação de imagens hiperespectrais aplicadas ao mapeamento de atividades agrícolas: abordagens com Deep Learning

Nesta seção, exploramos como técnicas de deep learning têm sido aplicadas à classificação de imagens hiperespectrais no contexto agrícola. Abordamos inicialmente as Redes Convolucionais com Autoatenção, destacando como modelos como SANet e HyperSFormer utilizam mecanismos de autoatenção para melhorar a precisão na classificação. Em seguida, discutimos as Otimizações e Modelos Híbridos em Redes Neurais Convolucionais, que combinam diferentes abordagens para aumentar a robustez e a eficácia dos modelos em cenários agrícolas complexos. Por fim, examinamos as Redes Convolucionais e Transformadores, onde frameworks como DMLP e HSI-TransUNet mostram-se eficazes na classificação de culturas e no mapeamento agrícola, oferecendo soluções para os desafios do setor.

3.2.1 Redes Convolucionais com Autoatenção

Em Zhang et al. (ZHANG et al., 2023), os autores propõem uma abordagem integrando múltiplos mecanismos de autoatenção e codificadores Transformer na rede SANet (A Self-Attention Network for Agricultural Hyperspectral Image Classification). Destacam a importância de considerar tanto informações espectrais quanto contextuais para melhorar a classificação de culturas. A ênfase na autocorrelação nas imagens hiperespectrais, combinada com a capacidade de capturar informações espaciais e espectrais complexas, ajuda a resolver desafios de mistura de características espaciais-espectrais nas fronteiras das culturas (YU et al., 2017; ZHANG et al., 2021). O modelo apresentou precisão geral acima de 95,8% nas bases testadas WHU-Hi. Apesar de sua eficácia, o SANet enfrenta limitações como generalização dos resultados e sensibilidade a hiperparâmetros, por isso é importante considerar essas possíveis limitações ao interpretar e aplicar os resultados do SANet em diferentes contextos agrícolas. Além disso, abordar essas limitações em futuras pesquisas pode contribuir para o aprimoramento e a robustez de métodos de classificação de imagens hiperespectrais agrícolas.

Xie et al. (XIE et al., 2023) desenvolvem uma nova abordagem de classificação de imagens hiperespectrais para agricultura, utilizando uma arquitetura de transformer. Este método, denominado HyperSFormer, substitui o codificador em SegFormer por um Swin Transformer aprimorado, mantendo o decodificador do SegFormer. Além disso, o estudo introduz o módulo de incorporação de hiper patch (HPE - Hyper Patch Embedding) para melhorar a extração de características espectrais e espaciais locais, e propõe um módulo de amostragem mínima log adaptativa (AMLS - Adaptive Min Log Sampling) para lidar com o problema de amostras insuficientes e desbalanceadas. Os resultados experimentais demonstram um desempenho robusto do HyperSFormer, superando métodos existentes em precisão e cobrindo dados de amostra desbalanceados e complexos. O modelo alcançou a acurácia global de 98,4%. Contudo, o trabalho menciona desafios persistentes relacionados a amostras limitadas e desequilibradas, o que pode afetar a generalização e a robustez do modelo de classificação de imagens hiperespectrais, apontando a necessidade de futuras otimizações para ampliar a aplicabilidade do modelo em diferentes cenários agrícolas (DEVLIN et al., 2018; LIANG et al., 2023).

Zhao et al. (ZHAO et al., 2023) exploram a classificação hiperespectral multitemporal de pastagens através de uma rede de transformadores. Este estudo inova ao aplicar o modelo MHCgT (Multi-Temporal Hyperspectral Classification using Transformer Network), que integra uma arquitetura hierárquica e mecanismo de autoatenção multi-cabeça, aprimorando a extração de características de séries temporais de dados hiperespectrais. Coletando 16.800 conjuntos de dados hiperespectrais multitemporais, o modelo alcançou uma precisão de 98,51%, superando métodos tradicionais como CNN, RNN-LSTM, SVM, RF e DT. Este desempenho destaca o potencial do MHCgT para identificação precisa de espécies em

gestão sustentável de pastagens e avaliação da diversidade de espécies. No entanto, o estudo enfrenta limitações, incluindo a interpretabilidade do modelo. Embora o modelo MHCgT demonstre precisão na classificação de pastagens, a interpretabilidade dos resultados e do processo de classificação pode não ter sido totalmente discutida, e que também não foram realizados testes em bases de dados públicas.

Em Hu et al. (HU et al., 2022a), apresenta-se uma rede de atenção tridimensional para classificação precisa de culturas usando imagens hiperespectrais de UAVs. O S3ANet (Spectral-Spatial-Scale Attention Network) utiliza módulos de atenção espectral, espacial e de escala para abordar a variabilidade espectral intraclasse e a heterogeneidade espacial. A abordagem melhora significativamente a precisão na classificação das culturas, obtendo mais de 96% de precisão geral em testes com menos de 50 pixels de treinamento por classe. Apesar dos resultados promissores, a validação externa do desempenho do S3ANet em diferentes conjuntos de dados ou em ambientes do mundo real pode ser uma limitação, pois a avaliação foi principalmente realizada nos conjuntos de dados específicos mencionados no artigo.

Em Zhang et al. (ZHANG et al., 2024), destaca-se a implementação de um dataset de referência denominado Heilongjiang (HLJ) para classificação de imagens hiperespectrais em larga escala, particularmente na classificação de culturas agrícolas. Composto por 319,685 e 318,942 amostras anotadas e 151 e 149 bandas espectrais, respectivamente, o dataset demonstrou robustez na diferenciação de culturas em cenários reais através de diferentes métodos de classificação, entre os quais o algoritmo 2D-Deform alcançou a melhor precisão geral (OA) de $97,54 \pm 0,10$, destacando-se como o mais eficaz. A pesquisa valida o dataset HLJ como uma ferramenta confiável para experiências de classificação, ilustrando o potencial do uso de dados hiperespectrais em larga escala para monitoramento agrícola e tomada de decisões informadas. Contudo, o estudo reconhece várias limitações. As principais desvantagens incluem a dependência de grandes quantidades de dados anotados, o que pode ser um desafio em cenários reais devido ao custo e ao esforço de coleta de dados. Além disso, a variação das condições de crescimento das culturas e a presença de múltiplas culturas intercaladas podem influenciar a precisão das classificações, exigindo futuras melhorias nos métodos de classificação para lidar com essas complexidades inerentes aos dados hiperespectrais.

3.2.2 Otimização e Modelos Híbridos em Redes Neurais Convolucionais

Hamza et al. (HAMZA et al., 2022) apresentam um modelo de classificação de culturas, o SSODTL-CC (Squirrel Search Optimization with Deep Transfer Learning-Enabled Crop Classification), que utiliza otimização por busca de esquilo (SSO - Squirrel Search Optimization) e aprendizado profundo transferido para identificar tipos de culturas em imagens hiperespectrais. Utilizando o MobileNet para extração de características e o

modelo BiLSTM (Bidirectional Long-Short Term Memory) para classificação, o SSODTL-CC alcançou resultados promissores em dois conjuntos de dados de referência (WHU-Hi-LongKou e WHU-Hi-HanChuan), com precisões de 99,23% e 97,15%. No entanto, o estudo destaca a eficácia do modelo SSODTL-CC para classificação de culturas em HSIs e sugere que o desempenho do modelo pode ser aprimorado no futuro com o design de modelos híbridos de Deep Learning.

Al Duhayyim et al. (DUHAYYIM et al., 2023) propuseram o modelo HOADTL-CC (Hurricane Optimization Algorithm with Deep Transfer Learning Driven Crop Classification) que combina otimização de algoritmo de furacão e aprendizado profundo para classificação de culturas em imagens hiperespectrais. Utilizando uma rede CapsNet para extração de características e uma rede neural Elman para classificação, o modelo foi avaliado em um conjunto de dados de referência, alcançando precisão de classificação (SHAHROUDNEJAD et al., 2018). O modelo atingiu a acurácia de 99,51%. No entanto, o estudo destaca a eficácia do modelo na classificação de culturas e sugere a aplicação de métodos de fusão baseados em DL para melhorar os resultados de classificação no futuro.

Em Lei et al. (LEI et al., 2021), um novo framework de classificação de uma única classe para mapeamento de culturas em larga escala usando dados de satélite multimodais é apresentado. O método DOCC (Deep One-Class Crop) utiliza aprendizado positivo e não rotulado para extrair automaticamente características importantes para a classificação de culturas usando dados de imagens hiperespectrais, multispectrais e SAR (Synthetic Aperture Radar). Este framework se destaca por requerer amostras da classe de interesse, simplificando o processo de rotulagem e reduzindo esforços redundantes associados à classificação multiclasse. O estudo mostra que o DOCC alcança precisão com 87,66% na classificação de diferentes tipos de culturas e demonstra robustez em diferentes modalidades de imagem. No entanto, há limitação de generalização, necessidade de ajustes de hiper parâmetros da CNN e da função de perda e dependência de amostras positivas representativas. Estas limitações apontam para a necessidade de futuras melhorias e adaptações do framework para diferentes contextos operacionais.

Yang e Hu (2022). (YANG; HU, 2022) investigam de que maneira as resoluções espacial e espectral afetam a exatidão na classificação de culturas agrícolas, empregando sistemas de sensoriamento remoto com veículos aéreos não tripulados. A pesquisa, ao utilizar o conjunto de dados WHU-Hi, indica que uma resolução espacial entre 0,5 e 1,0 metro é considerada otimizada, enquanto resoluções tanto elevadas quanto reduzidas são desnecessárias e podem comprometer a precisão dos resultados. A diminuição no número de bandas espectrais ocasionou uma leve alteração na acurácia da classificação, porém proporcionou uma considerável diminuição na carga computacional. Entre os métodos de aprendizado profundo analisados, a Rede Residual Espectral-Espacial (SSRN) evidenciou-se como a mais estável e precisa em diversas simulações de degradação, alcançando uma

acurácia global de 98%. Entretanto, a pesquisa apresenta restrições, como a exigência de um aprimoramento minucioso dos parâmetros de resolução, a fim de maximizar a exatidão, o que revela a necessidade de investigações adicionais para otimizar essas configurações em diversos contextos agrícolas.

Farmonov et al. (FARMONOV et al., 2023) apresentam um estudo abrangente sobre a classificação de tipos de culturas utilizando imagens hiperespectrais DESIS (Deutsches Zentrum für Luft- und Raumfahrt Earth Sensing Imaging Spectrometer) e algoritmos de aprendizado de máquina, destacando-se pela implementação de técnicas avançadas como Redes Neurais Convolucionais com atenção a wavelets (WA-CNN - Wavelet-Attention Convolutional Neural Network), Random forest e Máquinas de Vetores de Suporte (SVM). O método WA-CNN alcançou a maior precisão com uma acurácia global de 97,89%, beneficiando-se da transformação wavelet para extrair características relevantes e combiná-las com mecanismos de atenção espectral para aumentar a precisão na mapeamento de tipos de culturas. Contudo, o estudo enfrenta limitações como a dependência da resolução das imagens e a necessidade de vastas quantidades de dados para treinamento, o que pode restringir sua aplicabilidade em cenários com recursos limitados.

No trabalho de Wu et al. (WU et al., 2022), são propostos dois *frameworks* de classificação baseados em MLP (Multilayer Perceptrons) para imagens hiperespectrais de culturas: DMLP (Dilated MLP) e DMLPFFN (Dilation-Based MLP with Feature Fusion Network). O DMLP utiliza convolução dilatada para capturar contextos amplos, atingindo uma acurácia de 98,25% na base de dados Longkou, enquanto o DMLPFFN integra blocos residuais para uma extração de características mais profunda, apresentando uma acurácia de 99,16%. Testados em vários conjuntos de dados hiperespectrais, ambos os modelos demonstraram promissor desempenho, superando métodos tradicionais e de aprendizado profundo (YU et al., 2022). Entretanto, a capacidade dos modelos de generalizar para diferentes condições e cenários pode não ter sido totalmente avaliada. Esse desafio destacam área potencial para futuras melhorias nos modelos de classificação hiperespectral.

3.2.3 Redes Convolucionais e Transformadores

Em Wang et al. (WANG et al., 2022) propõem um método de classificação de imagens hiperespectrais, combinando um núcleo espacial espectral adaptativo com uma versão aprimorada do Vision Transformer (ViT - Vision Transformer). Este método introduz mecanismos de re-atenção para aumentar a diversidade dos mapas de atenção em diferentes níveis e um mecanismo local para utilizar informações locais e globais de forma eficaz, visando melhorar a precisão da classificação. O modelo atingiu como resultado a acurácia de 99,89% na base WHU-Hi-LongKou. Apesar dos resultados promissores em diversos conjuntos de dados, o estudo destaca a necessidade de futuras pesquisas para melhorar a arquitetura do transformador, desenvolver redes mais leves e reduzir a complexidade da

rede sem comprometer o desempenho.

Zhong et al.(ZHONG et al., 2020) desenvolvem uma metodologia avançada para classificação de culturas usando imagens hiperespectrais de alta resolução obtidas por UAVs. O modelo proposto, CNNCRF, combina redes CNN com campos aleatórios condicionais (CRF - Conditional Random Fields) para integrar características espectrais e espaciais locais, melhorando significativamente a precisão da classificação em áreas de pequena agricultura. Apesar de seu desempenho, com acurácia global acima de 93,70% nas bases testadas, validação e generalização são pontos a serem melhorados. A validação adequada do modelo em diferentes conjuntos de dados e a capacidade de generalização para novos casos são aspectos críticos. Se o modelo não for validado corretamente ou não puder generalizar para novos dados, sua utilidade prática pode ser limitada, destacando a necessidade de otimização contínua para aumentar sua aplicabilidade e eficiência em diferentes condições agrícolas.

Niu et al.(NIU et al., 2022), desenvolvem um modelo de segmentação semântica chamado HSI-TransUNet (Hyperspectral Transformer U-Net), destinado ao mapeamento de culturas usando imagens hiperespectrais de UAVs. Este modelo combina atenção espectral e camadas de Transformer para explorar as informações espaciais e espectrais, superando modelos anteriores em termos de precisão na classificação de culturas. Apesar do seu desempenho, com uma acurácia global de 86,05%, o HSI-TransUNet enfrenta desafios, como a validação do modelo foi realizada em um conjunto de dados específico, e não há menção de validação em diferentes conjuntos de dados ou em diferentes condições de captura de imagens. Isso pode limitar a compreensão da robustez do modelo em diferentes cenários. Estas limitações indicam a necessidade de otimização futura para ampliar sua aplicabilidade e eficiência.

No estudo realizado por Patel et al. (PATEL et al., 2023), é analisada a categorização de diferentes tipos de culturas por meio de imagens hiperespectrais, utilizando uma metodologia de aprendizado profundo fundamentada na transferência de aprendizado. O estudo investiga a transferência de aprendizado em suas vertentes heterogênea e homogênea, empregando modelos pré-treinados, como VGG16, VGG19, ResNet e DenseNet, além de redes neurais convolucionais personalizadas, como 2DCNN e 3DCNN. Os resultados demonstram-se encorajadores, uma vez que o modelo DenseNet obteve uma precisão que ultrapassa 99% nos conjuntos de dados de Indian Pines, Pavia University e Salinas, utilizando apenas 15% das amostras rotuladas. Isso ressalta a eficácia das abordagens de transferência de aprendizado em cenários com restrita disponibilidade de dados devidamente rotulados. No entanto, uma limitação relevante indicada é a perda de informação espectral ocasionada pela aplicação de PCA, a fim de adequar a dimensionalidade das imagens hiperespectrais à entrada exigida pelos modelos previamente treinados, o que pode restringir a total exploração das valiosas informações espectrais contidas nas imagens hiperespectrais.

Ademais, a adaptação de modelos para múltiplos canais em imagens hiperespectrais enfrenta desafios consideráveis, evidenciando a necessidade de modificações nos modelos a fim de maximizar a exploração da riqueza espectral presente nos dados de imagens hiperespectrais.

Vaishnavi et al. (VAISHNAVI et al., 2022), a pesquisa se concentra no desenvolvimento de um modelo 2D-CNN para a classificação eficaz de imagens hiperespectrais em aplicações agrícolas. Este modelo é especialmente otimizado para capturar e analisar características espectrais e espaciais de culturas, mostrando resultados promissores em conjuntos de dados reconhecidos como *WHU-HI*, *Indian Pines*, *University of Pavia* e *Salinas*. A abordagem 2D-CNN proporciona uma análise detalhada das interações espectrais e espaciais das imagens, resultando em uma classificação precisa das culturas. O resultado foi com a base University of Pavia que obteve uma acurácia global de 99,93%. Contudo, o desempenho do modelo pode variar significativamente quando aplicado a novos conjuntos de dados ou diferentes condições ambientais. A necessidade de desenvolver e testar o modelo em uma ampla variedade de condições e tipos de dados é essencial para garantir a robustez e a aplicabilidade generalizada do método desenvolvido.

Galodha et al. (GALODHA et al., 2022b) desenvolveram um modelo avançado usando CNN para a classificação de culturas a partir de imagens hiperespectrais capturadas por drones. O modelo utiliza arquiteturas Conv-4 e Conv-6, variando os tamanhos de kernel para extrair múltiplas características das imagens, demonstrando que o Conv-6 proporciona precisão, atingindo 87,80% de acurácia global. Os resultados incluem medidas visuais e quantitativas que validam a superioridade do CNN Conv-6 sobre o Conv-4, e são acompanhados de gráficos representando a precisão de treinamento e validação em função do número de iterações, o que ajuda a determinar o número de épocas necessárias para uma classificação e previsão eficazes. Contudo, o modelo enfrenta desafios como a uniformidade das imagens hiperespectrais e a necessidade de ajustes finos nos parâmetros para maximizar o desempenho, apontando para a necessidade de otimizações futuras para aumentar a aplicabilidade do modelo em diferentes cenários agrícolas. É válido destacar também a necessidade de testes em bases de dados públicas com mais métricas de avaliação.

Bhosle e Musande. (BHOSLE; MUSANDE, 2019), exploram o uso de redes neurais convolucionais (CNNs) para a classificação de uso e cobertura do solo (LULC - Land Use Land Cover) e identificação de culturas utilizando imagens hiperespectrais. A pesquisa emprega o conjunto de dados Indian Pines e dados de uma área de estudo específica, usando um sensor EO-1 Hyperion. O estudo demonstra que ajustes em parâmetros como taxa de aprendizado, tamanho do filtro e funções de ativação podem significativamente melhorar a precisão das CNNs, alcançando 97% de precisão no conjunto de dados Indian Pines e 78% na área de estudo específica. Apesar dos resultados promissores, o estudo reconhece a necessidade de uma seleção criteriosa de parâmetros e mais pesquisas para

otimizar a aplicabilidade das CNNs em diversas condições de dados hiperespectrais.

Agilandeewari et al.([AGILANDEESWARI et al., 2022](#)) discutem a classificação de culturas usando imagens hiperespectrais obtidas por UAV, enfocando a aplicação de uma rede neural convolucional 2D para selecionar bandas eficazes. Inicialmente, as bandas são particionadas e quantizadas com base em métricas como entropia e índices de vegetação. O modelo CNN aprimora a precisão da classificação em culturas distintas nos datasets de Indian Pines e Salinas, alcançando precisão geral de 97,62% e 96,08% respectivamente. No entanto, o estudo revela limitações na capacidade do modelo de CNN 2D em integrar e explorar profundamente as informações espectrais, que são críticas para a identificação detalhada das características das culturas. Essa constatação sublinha a necessidade de desenvolvimento futuro, seja por meio de modelos híbridos que incorporem capacidades de análise espectral mais sofisticadas ou pelo aprimoramento das arquiteturas de CNN para abranger análises espectrais mais detalhadas.

Galodha et al.([GALODHA et al., 2022a](#)) realizam uma investigação sobre a utilização de Redes Neurais Convolucionais e Redes Residuais na classificação espectral-espacial de imagens hiperespectrais capturadas por drones e plataformas terrestres. A pesquisa aplica uma estrutura profunda de CNN composta por seis camadas, otimizando pesos em diferentes fases do procedimento de classificação, que abrange convolução, pooling e uma camada densa. Essa metodologia atingiu uma precisão excepcional, apresentando uma acurácia global de 97,16% ao empregar o otimizador RMSProp em dados de drones, com acurácias específicas para culturas variando de 80,566% a 89,25%. Não obstante os índices de precisão, a pesquisa destaca desafios substanciais, tais como a dependência de condições ideais para a captura de imagens e a variação no desempenho dos otimizadores em diferentes configurações. A intrínseca natureza da arquitetura influencia diretamente o tempo de processamento, configurando um compromisso entre precisão e eficiência operacional. Os resultados destacam a crescente relevância do aprendizado profundo na agricultura de precisão, indicando que aprimoramentos nas metodologias de pré-processamento e validação cruzada podem aumentar significativamente a eficácia dessas tecnologias avançadas.

No estudo realizado por Bhosle e Musande ([BHOSLE; MUSANDE, 2022](#)), é realizada uma avaliação do desempenho dos modelos de Redes Neurais Convolucionais, em comparação com Autoencoders Convolucionais e Redes Neurais Profundas, voltada para a classificação de culturas em imagens hiperespectrais. A pesquisa revela, ao utilizar os conjuntos de dados Indian Pines e informações de uma área de estudo específica, que a Rede Neural Convolucional otimizada obtém um desempenho superior em comparação a outros métodos, atingindo uma taxa de precisão global de 97% no conjunto de dados Indian Pines e de 78% na área de estudo. Os resultados apontam para a eficácia da Rede Neural Convolucional na extração de características e na classificação precisa, mesmo diante de uma quantidade limitada de amostras de treinamento. Entretanto, a pesquisa

aponta restrições, como a sensibilidade do modelo em relação a alterações nos parâmetros de treinamento e a exigência de otimização constante para enfrentar a variabilidade dos dados provenientes do sensoriamento remoto. Ademais, a execução em áreas reduzidas de investigação indica a importância de confirmar a transferibilidade do modelo para variados contextos e escalas, o que é um aspecto fundamental para a aplicação prática na agricultura de precisão.

Por fim, a discriminação de diferentes tipos de culturas foi discutida por Singh et al. (SINGH et al., 2022), utilizando a técnica de extração de endmembers Geo-Stat (GSEE - Geo-Stat Endmember Extraction) em conjunto com algoritmos de aprendizado de máquina aplicados a imagens hiperespectrais. A pesquisa utilizou métodos avançados de classificação supervisionada, tais como SVM, Mapeador de Ângulo Espectral (SAM - Spectral Angle Mapper) e CNN, com o objetivo de identificar nove variedades de culturas. Os endmembers foram obtidos por meio do algoritmo GSEE, o qual possibilitou a identificação de pixels puros e a criação de uma biblioteca espectral que fundamentou a classificação das culturas. Esse método evidenciou eficácia, com o classificador CNN atingindo uma acurácia global de 89,065% e um coeficiente kappa de 0,871. Apesar dos resultados encorajadores, a pesquisa admite restrições, como a dependência de abrangentes conjuntos de dados de treinamento e a necessidade de uma criteriosa seleção dos endmembers, visando evitar a inclusão de pixels mistos, os quais podem comprometer a exatidão da classificação. Ademais, o desafio de ajustar esses métodos sofisticados para que sejam viáveis em diversos contextos e tipos de culturas permanece, evidenciando uma área de pesquisa relevante para investigações futuras.

Os estudos analisados na Tabela 7 sobre Deep Learning proporcionaram uma fundamentação robusta para a elaboração e aprimoramento do modelo sugerido nesta dissertação. A avaliação de métodos como Redes de Autoatenção (SAnet), Transformadores de Visão (ViT) e CNN evidenciou como técnicas sofisticadas podem alcançar bons níveis de precisão em dados hiperespectrais, com alguns estudos reportando até 99,89% de precisão. Essas pesquisas, que investigam desde as redes convolucionais tradicionais até estruturas mais intrincadas como Transformers e Redes Residuais (ResNet), tiveram um impacto direto no desenvolvimento da minha abordagem híbrida de CNN. Especificamente, a incorporação de métodos de fusão espectral-espacial e a implementação de redes profundas possibilitaram uma melhor compreensão das propriedades multidimensionais dos dados hiperespectrais, levando a uma classificação mais acurada de zonas agrícolas. Portanto, os avanços observados nos artigos com Deep Learning ajudaram a definir as direções metodológicas para o aprimoramento do modelo proposto, garantindo um desempenho robusto e eficiente na classificação e mapeamento hiperespectral.

Tabela 7 – Resumo dos artigos com aprendizado de máquina com Deep Learning

Artigo	Aplicação Investigada	Técnica Utilizada	Resultados	Bases de Dados
Zhang et al. (ZHANG et al., 2023)	Classificação de culturas hiperespectrais usando SANet	SANet (Self-Attention Network) integrando codificadores Transformer	Precisão geral acima de 95,8% nas bases WHU-Hi	WHU-Hi-HongHu, WHU-Hi-HanChuan, WHU-Hi-LongKou
Xie et al. (XIE et al., 2023)	Classificação de culturas hiperespectrais para agricultura	HyperFormer utilizando Swin Transformer e módulos de atenção	Acurácia global de 98,4%	Indian Pines, HanChuan, HongHu
Zhao et al. (ZHAO et al., 2023)	Classificação hiperespectral multitemporal de pastagens	MHCgT (Multi-Temporal Hyperspectral Classification using Transformer Network)	Precisão de 98,51%	Região Autônoma da Mongólia Interior, China
Hu et al. (HU et al., 2022a)	Classificação precisa de culturas usando imagens hiperespectrais de UAVs	S3ANet (Spectral-Spatial-Scale Attention Network)	Mais de 96% de precisão geral	WHU-Hi-HongHu, WHU-Hi-HanChuan e WHU-Hi-JiaYu
Zhang et al. (ZHANG et al., 2024)	Classificação de imagens hiperespectrais em larga escala com o dataset Heilongjiang (HLJ)	2D-Deform	Acurácia geral (OA) de $97,54 \pm 0,10$	HLJ-Raohe dataset, HLJ-Yan, LongKou dataset, HanChuan, RiverEstuary, Indian-Pines, Salinas
Hamza et al. (HAMZA et al., 2022)	Classificação de culturas hiperespectrais	SSODTL-CC (Squirrel Search Optimization com Deep Transfer Learning)	99,23% na base WHU-Hi-LongKou e 97,15% na base WHU-Hi-HanChuan	WHU-Hi-LongKou, WHU-Hi-HanChuan
Al Duhayyim et al. (DUHAYYIM et al., 2023)	Classificação de culturas hiperespectrais	HOADTL-CC (Hurricane Optimization Algorithm com Deep Transfer Learning)	Acurácia de 99,51%	WHU-Hi-LongKou
Lei et al. (LEI et al., 2021)	Mapeamento de culturas em larga escala usando dados multi-modais	DOCC (Deep One-Class Crop)	Acurácia de 87,66%	Zhuhai-1
Yang e Hu. (YANG; HU, 2022)	Classificação de culturas agrícolas usando UAVs	SSRN (Spectral-Spatial Residual Network)	Acurácia geral (OA) de 98%	WHU-Hi-LongKou, WHU-Hi-HanChuan, WHU-Hi-HongHu
Farmonov et al. (FARMONOV et al., 2023)	Classificação de tipos de culturas usando imagens hiperespectrais DESIS	WA-CNN (Wavelet-Attention Convolutional Neural Network)	Acurácia global de 97,89%	Região de Mezőhegyes, na Hungria, University of Pavia, Indian Pines, WHU-Hi-LongKu
Wu et al. (WU et al., 2022)	Classificação de culturas hiperespectrais	DMLP (Dilated MLP) e DMLPFFN (Dilation-Based MLP with Feature Fusion Network)	98,25% (DMLP) e 99,16% (DMLPFFN)	Salinas, KSC, LongKou, HanChuan
Wang et al. (WANG et al., 2022)	Classificação de imagens hiperespectrais	Vision Transformer (ViT) com núcleo espacial espectral adaptativo	Acurácia de 99,89% na base WHU-Hi-LongKou	Indian Pines, Pavia University, Xuzhou, LongKou
Zhong et al. (ZHONG et al., 2020)	Classificação de culturas usando imagens hiperespectrais de UAVs	CNNCRF (CNN com CRF)	Acurácia global acima de 93,70%	WHU-Hi-HongHu, WHU-Hi-HanChuan, WHU-Hi-LongKou
Niu et al. (NIU et al., 2022)	Mapeamento de culturas usando UAVs	HSI-TransUNet (Hyperspectral Transformer U-Net)	Acurácia global de 86,05%	Shenzhou City, na província de Hebei, China
Patel et al. (PATEL et al., 2023)	Classificação de culturas usando transferência de aprendizado	VGG16, VGG19, ResNet, DenseNet, 2DCNN, 3DCNN	Precisão superior a 99%	Indian Pines, Salinas, University of Pavia
Vaishnavi et al. (VAISHNAVI et al., 2022)	Classificação de culturas usando imagens hiperespectrais	2D-CNN otimizado para imagens hiperespectrais	Acurácia global de 99,93% na base University of Pavia	University of Pavia, Indian Pines, Salinas Scene, WHU-Hi-LongKou, WHU-Hi-HanChuan, WHU-Hi-HongHu
Galodha et al. (GALODHA et al., 2022b)	Classificação de culturas usando imagens hiperespectrais de drones	CNN (Conv-4 e Conv-6)	Acurácia global de 87,80%	Região de Gandhi Krishi Vignan Kendra (GKVK), localizada ao norte de Bangalore, na Índia.
Bhosle e Musande (BHOSLE; MUSANDE, 2019)	Classificação de culturas usando imagens hiperespectrais	CNN otimizada para Indian Pines	97% de precisão	Indian Pines
Agilandeewari et al. (AGILANDEESWARI et al., 2022)	Classificação de culturas usando imagens hiperespectrais obtidas por UAV	2D-CNN para seleção de bandas	Precisão geral de 97,62% (Indian Pines) e 96,08% (Salinas)	Indian Pines, Salinas
Galodha et al. (GALODHA et al., 2022a)	Classificação espectral-espacial de imagens hiperespectrais	CNN e ResNet com RMS-Prop	Acurácia global de 97,16%	Cubert FirefEYE SE S-185 Hyperspectral Image Dataset (Bangalore, Índia)
Bhosle e Musande (BHOSLE; MUSANDE, 2022)	Classificação de culturas em imagens hiperespectrais	CNN, Autoencoders Convolutionais e DNN	Precisão geral de 97% (Indian Pines) e 78% (área de estudo específica)	Indian Pines, Aurangabad e Maharashtra State, India
Singh et al. (SINGH et al., 2022)	Classificação de culturas em imagens hiperespectrais	GSEE, SVM, SAM, CNN (AVIRIS-NG)	Acurácia global de 89,065%, coeficiente kappa de 0,871	AVIRIS-NG Anand Dataset (Gujarat, Índia)

Os estudos analisados utilizam uma variedade de bases de dados, desde conjuntos amplamente conhecidos, como Indian Pines e Salinas, até dados regionais específicos, como o HLJ-Raohe dataset (China) e o AVIRIS-NG Anand Dataset (Índia). Essa diversidade de fontes evidencia a aplicabilidade das técnicas de classificação hiperespectral em diferentes contextos geográficos e agrícolas. No entanto, ela também revela um desafio crítico: a dificuldade de generalização dos modelos para diferentes regiões e tipos de culturas. Por exemplo, enquanto alguns estudos alcançam acurácias superiores a 99% em bases de dados específicas, outros, como o trabalho de Lei et al., obtêm resultados mais modestos 87,66% ao lidar com mapeamento de culturas em larga escala. Essa discrepância sugere que a variabilidade e a complexidade dos dados podem comprometer significativamente o desempenho dos modelos em cenários reais. Além disso, questões como a complexidade computacional e a necessidade de modelos mais generalizáveis ainda precisam ser abordadas para que essas tecnologias possam ser amplamente adotadas na prática agrícola. Soluções que integrem técnicas de transferência de aprendizado ou adaptação de modelos para diferentes contextos podem ser caminhos promissores para superar esses desafios.

3.3 Considerações Finais

O estudo conduzido revelou que as CNNs, especialmente as arquiteturas 2D e 3D, demonstram boa eficiência em bases de dados públicas, o que guiou a seleção metodológica deste estudo. Ao realizar a pesquisa bibliográfica, observou-se uma variedade de métodos nas pesquisas já realizadas, contudo, várias limitações e lacunas foram identificadas. Dentre os problemas identificados, estão a ausência de testes em bases de dados públicas amplamente reconhecidas, a utilização limitada de métricas de avaliação, a dependência de grandes volumes de dados de treinamento e a suscetibilidade a alterações nos parâmetros de treino. Adicionalmente, notou-se a necessidade de aprimoramento constante para gerir a variabilidade dos dados de sensoriamento remoto, além da ausência de validação cruzada e avaliação da adaptabilidade dos modelos a variados contextos e escalas. Foi também identificado o desafio das CNNs 2D em integrar e analisar profundamente as informações espectrais, cruciais para a identificação minuciosa das propriedades das culturas. Além disso, a aplicação restrita dos modelos em variados cenários agrícolas e a ausência de clareza sobre o pré-processamento dos dados também foram identificadas. Esses elementos muitas vezes prejudicam a aplicabilidade prática dos modelos em diversas culturas e em diferentes condições ambientais.

Este estudo sugeriu a criação de uma estratégia híbrida, empregando redes neurais com arquiteturas CNN 2D e 3D, testada em oito bases de dados públicas e avaliada com sete métricas, incluindo a validação cruzada. Esta opção teve como objetivo minimizar as restrições identificadas, reforçar a solidez e a aplicabilidade dos modelos de classificação agrícola, além de expandir sua capacidade de uso prático. As aplicações potenciais incluem

o acompanhamento em tempo real de colheitas, a identificação antecipada de enfermidades nas plantas, a otimização do uso de recursos hídricos e insumos agrícolas, bem como o planejamento e administração sustentável do solo. Portanto, a expectativa é que a metodologia sugerida tenha um impacto significativo no progresso das tecnologias de sensoriamento remoto utilizadas na agricultura.

4 Materiais e Métodos

Este capítulo detalha os materiais e métodos utilizados no desenvolvimento do modelo proposto, incluindo as bases de dados selecionadas, as técnicas de pré-processamento aplicadas, a arquitetura de CNN empregada para a classificação das imagens, e as métricas de avaliação usadas para validar e comparar o modelo proposto com outros métodos da literatura.

4.1 Bases de Dados Utilizadas

Foram selecionadas bases de dados hiperespectrais amplamente utilizadas na literatura para o mapeamento de áreas agrícolas.

Tabela 8 – Bases de Dados.

Base de Dados	Dimensão	Imagens	Classes	Bandas
Indian Pines (PATEL et al., 2023)	145×145	10.249	16	200
Salinas (PATEL et al., 2023)	512×217	54.129	16	204
Universidade de Pavia (PATEL et al., 2023)	610×340	42.776	9	103
KSC (BASCO, 2021)	512×614	5.211	13	176
Botswana (BASCO, 2021)	1476×256	3.248	14	145
WHU-Hi-LongKou (VAISHNAVI et al., 2022)	550×400	204.542	9	270
WHU-Hi-HanChuan (VAISHNAVI et al., 2022)	1217×303	257.530	16	274
WHU-Hi-HongHu (VAISHNAVI et al., 2022)	940×475	386.693	22	270

A Tabela 8 resume o conjunto abrangente de bases de dados hiperespectrais utilizadas para a análise e validação do modelo de classificação proposto. Entre as bases de dados incluídas estão: Indian Pines, com 10.249 imagens distribuídas em 16 classes e 200 bandas espectrais; Salinas, com 54.129 imagens e 204 bandas; e Universidade de Pavia, que apresenta 42.776 imagens, 9 classes e 103 bandas. Além dessas, a tabela também enumera os conjuntos de dados de KSC e Botswana, respectivamente com 5.211 e 3.248 imagens, abrangendo 13 e 14 classes com 176 e 145 bandas espectrais. As bases de dados WHU-Hi, incluindo LongKou, HanChuan e HongHu, representam os conjuntos mais volumosos, com centenas de milhares de imagens e mais de 270 bandas cada, oferecendo uma diversidade significativa em termos de classes de objetos, que vão de 9 a 22. Este levantamento permite uma compreensão das características e variações intrínsecas aos diferentes cenários e tipos de vegetação, essenciais para o avanço das técnicas de classificação em aplicações hiperespectrais. A riqueza e variedade das bases de dados selecionadas asseguram uma robusta validação e generalização do modelo desenvolvido, aumentando a relevância e aplicabilidade dos resultados da pesquisa.

4.2 Modelo Proposto

O modelo criado para a classificação de imagens hiperespectrais foi concebido com a finalidade de explorar os benefícios das arquiteturas convolucionais tridimensionais e bidimensionais, aliado a técnicas avançadas, como a Normalização em Lote (Batch Normalization) e as convoluções profundas Depthwise. Este modelo tem como objetivo registrar, de forma eficaz, tanto as propriedades espaciais quanto as espectrais das imagens, garantindo o desempenho na atividade de classificação.

Na Figura 12 é apresentada a representação do modelo proposto. Inicialmente, a base de imagens hiperespectrais passa por uma redução de dimensionalidade utilizando o PCA. Em seguida, são criados os patches 3D, que são então processados pelo modelo proposto.



Figura 12 – Representação do modelo proposto neste trabalho.

Na etapa de pré-processamento, o cubo de dados hiperespectrais espectrais-espaciais é representado por $A \in \mathbb{R}^{M \times N \times D}$, sendo A a matriz original, M a largura, N a altura e D a quantidade de bandas espectrais ou profundidade. Cada pixel HSI presente em A contém D medidas espectrais e configura um vetor rotulado one-hot $Y = (y_1, y_2, \dots, y_C) \in \mathbb{R}^{1 \times 1 \times C}$, em que C indica as categorias de cobertura do solo.

No entanto, os pixels hiperespectrais em A exibem classes de cobertura do solo mistas, introduzindo alta variabilidade intraclasse e similaridade interclasse em A . É um grande desafio para qualquer modelo lidar com esse problema. Para remover primeiro a redundância espectral, a análise de componentes principais (PCA) tradicional é aplicada sobre os dados HSI originais (A) ao longo das bandas espectrais.

A utilização do PCA no pré-processamento dos dados hiperespectrais visa principalmente a redução de dimensionalidade e a eliminação de redundâncias nas bandas espectrais. As imagens hiperespectrais contêm muitas bandas que, frequentemente, apresentam alta correlação entre si, o que pode aumentar a complexidade computacional e o risco de overfitting no modelo. Ao aplicar o PCA, é possível reduzir a quantidade de dados mantendo a maior parte da variabilidade informativa, o que melhora a eficiência do treinamento da rede neural convolucional. Além disso, a redução de dimensionalidade acelera a convergência do modelo e melhora a generalização, tornando o processo mais robusto e eficiente. Na Figura 12 o PCA reduz o número de bandas espectrais de D para B enquanto mantém as mesmas dimensões espaciais ou seja, largura M e altura

N . Reduzimos apenas as bandas espectrais de modo a preservar a informação espacial, que é muito importante para o reconhecimento de qualquer objeto. A escolha de $B = 3$ componentes principais foi feita com base na análise da variância explicada.

Representamos o cubo de dados reduzido pelo PCA por $X \in \mathbb{R}^{M \times N \times B}$, onde X é a entrada modificada após o PCA, M é a largura, N é a altura, e B é o número de bandas espectrais após o PCA.

Antes de utilizar o modelo proposto na classificação de imagens hiperespectrais, após a aplicação do PCA, os cubos de dados HSI são divididos em pequenos patches 3D sobrepostos, cujos rótulos verdadeiros são determinados pelo rótulo do pixel central. A criação desses patches tem como objetivo preservar e explorar tanto as informações espectrais quanto espaciais, permitindo que o modelo capture relações contextuais entre os pixels. Essa abordagem é essencial para a utilização de redes neurais convolucionais 3D (3D-CNN), que extraem características espectro-espaciais, melhorando a diferenciação entre classes e tornando a classificação mais robusta contra ruídos. Criamos os patches 3D vizinhos $P \in \mathbb{R}^{R \times R \times B}$ a partir de X , centralizados na localização espacial (α, β) , cobrindo a janela de $R \times R$ ou a extensão espacial e todas as B bandas espectrais. A quantidade total de patches 3D produzidos (n) a partir de X é expressa pela fórmula $(M - R + 1) \times (N - R + 1)$. Desta forma, o patch 3D situado na coordenada (α, β) , representado por $P_{\alpha, \beta}$, abrange a largura de $\alpha - (R - 1)/2$ a $\alpha + (R - 1)/2$, a altura de $\beta - (R - 1)/2$ a $\beta + (R - 1)/2$, além de todas as B bandas espectrais do cubo de dados limitados pela Análise de Componentes Principais X .

A primeira camada do modelo é uma camada convolucional tridimensional (Conv3D), que aplica 8 filtros tridimensionais com tamanho de kernel $(3, 3, 3)$ sobre o tensor de entrada. Esta camada utiliza a função de ativação ReLU, que é amplamente utilizada para introduzir não-linearidade na rede. Para garantir que as dimensões da saída permaneçam as mesmas da entrada, foi utilizado padding do tipo same. Além disso, a camada é seguida por uma normalização em lote (BatchNormalization), o que contribui para estabilizar o processo de treinamento ao normalizar a saída da convolução.

Em seguida, uma segunda camada convolucional tridimensional é aplicada, desta vez com 16 filtros e um kernel de tamanho $(3, 3, 1)$. Esta camada não utiliza padding, resultando em uma leve redução das dimensões espaciais da saída. Novamente, Batch Normalization é empregado para melhorar a estabilidade do treinamento e acelerar a convergência.

Após as operações convolucionais tridimensionais, a saída da segunda camada 3D é remodelada através de um redimensionamento. Essa etapa é necessária para transformar as dimensões da saída em uma forma que seja compatível com a aplicação de uma convolução bidimensional (2D). Mais especificamente, a saída é ajustada para combinar as últimas duas dimensões, permitindo a utilização subsequente da camada convolucional 2D.

Em seguida, realiza-se a aplicação de uma camada de convolução profunda (Depthwise) utilizando um kernel de dimensão (3, 3). Essa camada executa a convolução de maneira independente em cada canal de entrada, o que diminui de forma significativa a quantidade de parâmetros do modelo. A camada de convolução profunda é sucedida por uma convolução ponto a ponto em 2D, utilizando um kernel de (1, 1), que são aplicados 64 filtros para integrar os canais derivados da convolução profunda. A combinação de DepthwiseConv2D e Convolução 2D representa uma abordagem eficaz para diminuir a complexidade computacional, preservando a habilidade do modelo em identificar características importantes.

A saída da última camada convolucional é achatada através de uma camada Flatten, transformando o tensor em um vetor unidimensional que pode ser processado por camadas densas totalmente conectadas. Duas camadas densas são, então, aplicadas, com 256 e 128 unidades, respectivamente, ambas utilizando a função de ativação ReLU. Para mitigar o risco de sobreajuste (overfitting), uma camada de Dropout com taxa de 30% é inserida após cada uma das camadas densas.

Finalmente, a camada de saída do modelo é composta por C unidades, correspondendo ao número de classes do conjunto de dados. A função de ativação softmax é utilizada nesta camada, gerando as probabilidades de pertencimento a cada classe para as amostras de entrada.

O modelo foi reinicializado a cada nova rodada de treino para garantir que não houvesse influência de treinos anteriores. O processo de ajuste foi feito usando o método Adam, que ajuda a encontrar o melhor caminho de aprendizado com uma taxa de 0,001. Para medir o desempenho do modelo, foi utilizada a acurácia, que indica a porcentagem de acertos na classificação.

Para evitar que o modelo se ajustasse demais aos dados e perdesse capacidade de generalização, foi usado um recurso que interrompe o treinamento automaticamente após 10 rodadas sem melhorias, garantindo que o melhor desempenho fosse mantido.

Este modelo foi desenvolvido para equilibrar a complexidade computacional com a capacidade de generalização, aproveitando tanto a riqueza da convolução tridimensional para capturar as características espectrais quanto a eficiência da convolução bidimensional para as características espaciais.

Na Tabela 9 apresenta-se a arquitetura detalhada do modelo proposto, incluindo as camadas utilizadas, os formatos de saída (Output Shape) de cada camada, e o número de parâmetros associados (Param #). O modelo começa com uma camada de entrada (InputLayer) que recebe dados com a forma (25, 25, 3, 1), onde 25x25 são as dimensões espaciais dos patches, 3 representa as bandas espectrais após a aplicação de PCA, e 1 indica que os dados são monocanais.

Camada	Forma de Saída	Param #
Camada de Entrada	(25, 25, 3, 1)	0
Convolução 3D	(25, 25, 3, 8)	224
Normalização em Batch	(25, 25, 3, 8)	32
Convolução 3D	(23, 23, 3, 16)	1.168
Normalização em Batch	(23, 23, 3, 16)	64
Redimensionamento	(23, 23, 48)	0
Convolução Depthwise 2D	(21, 21, 48)	480
Convolução 2D	(21, 21, 64)	3.136
Camada Densa	(256)	7.225.600
Dropout	(256)	0
Camada Densa	(128)	32.896
Dropout	(128)	0
Camada de Saída	(16)	2.064
Total		7.265.664

Tabela 9 – Resumo da arquitetura proposta com um tamanho de janela de 25×25 . A última camada é baseada no conjunto de dados Indian Pines.

O número total de parâmetros do modelo é 7.265.664, dos quais 7.265.616 são treináveis, ocupando um total de 27,72 MB. Esses valores refletem a complexidade e a capacidade do modelo proposto de capturar informações detalhadas de imagens hiperespectrais.

4.3 Divisão dos Dados e Treinamento

Com o objetivo de analisar a eficácia do modelo, os dados foram inicialmente divididos em 70% para treinamento e 30% para teste, garantindo que os conjuntos de dados de treinamento e teste não se misturassem. Dentro dos 70% destinados ao treinamento, foi aplicada a técnica de validação cruzada estratificada utilizando K-Fold com cinco partições. Essa abordagem assegura que cada subdivisão mantenha a mesma proporção de classes presentes no conjunto original, proporcionando uma distribuição equilibrada das classes. Em cada iteração, uma partição foi utilizada para validação, enquanto as outras quatro foram usadas para treinamento. Esse processo foi repetido para todos os cinco folds, resultando em cinco treinamentos independentes. A escolha da validação cruzada K-Fold visa fornecer uma avaliação mais robusta do desempenho do modelo, já que ele é treinado e validado em diferentes subconjuntos dos dados, o que ajuda a reduzir o risco de overfitting e a melhorar a generalização do modelo.

Para cada fold, o modelo foi recompilado e treinado com os dados correspondentes. O modelo utilizado foi uma rede neural convolucional, com otimização baseada no Adam (com taxa de aprendizado de 0.001) e utilizando uma função de perda, adequada para problemas de classificação multiclasse. A rede foi treinada por até 20 épocas em cada fold, com um tamanho de lote de 128. Durante o treinamento, a EarlyStopping foi aplicada

para evitar overfitting, monitorando a perda de validação e interrompendo o treinamento caso não houvesse melhorias após 10 épocas consecutivas.

Além disso, o tempo de treinamento de cada fold foi registrado para avaliar a eficiência do processo. O desempenho do modelo foi mensurado por meio da acurácia e perda, sendo calculados após cada fold.

Os resultados finais foram obtidos pela média das métricas de cada fold, incluindo a acurácia média, a perda média e o tempo médio de treinamento por fold. Essas métricas oferecem uma avaliação abrangente da eficácia e eficiência do modelo treinado.

Esta divisão foi implementada utilizando a técnica de validação cruzada estratificada com cinco divisões (Stratified K-Fold), por ser uma abordagem amplamente usada em análise estatística e aprendizado de máquina para assegurar que cada divisão seja representativa do conjunto de dados completo. Optamos por usar cinco divisões por se tratar de um equilíbrio entre robustez e eficiência. Cinco divisões proporcionam uma boa quantidade de repetições para avaliar o desempenho do modelo em diferentes amostras de dados, ao mesmo tempo em que evitam um custo computacional excessivo, que seria observado em validações com um número muito maior de divisões. Essa técnica é particularmente útil em cenários onde o balanceamento das classes é importante para a acurácia do modelo, onde cada conjunto de dados manteve a proporção original das classes para assegurar que todos os segmentos do dataset fossem representativos e consistentes. O formato final dos dados para entrada no modelo incluiu redimensionamento para acomodar as dimensões espaciais e espectrais apropriadas, e a conversão dos rótulos para formato categórico, facilitando a aplicação das técnicas de classificação multiclasse usadas no treinamento do modelo.

4.4 Métricas de Avaliação Utilizadas

Para a análise da eficácia do modelo, foram consideradas sete métricas de desempenho. Cada uma dessas análises fornece uma visão distinta sobre o desempenho do modelo, abordando vários elementos de exatidão e consistência. A base para o cálculo de muitas dessas análises é uma matriz de confusão. Para um problema de classificação com C classes, a matriz de confusão M é uma matriz $C \times C$ onde:

A matriz de confusão constitui a fundamentação para a regulamentação de diversas considerações. Em um contexto de classificação com C categorias, a matriz de confusão M se apresenta como uma matriz de dimensão $C \times C$, na qual:

$$M = \begin{bmatrix} M_{1,1} & M_{1,2} & \dots & M_{1,C} \\ M_{2,1} & M_{2,2} & \dots & M_{2,C} \\ \vdots & \vdots & \ddots & \vdots \\ M_{C,1} & M_{C,2} & \dots & M_{C,C} \end{bmatrix}$$

Aqui, $M_{i,j}$ representa o número de instâncias que pertencem à classe C_i (classe verdadeira) e foram classificadas como C_j (classe prevista).

Verificação de Casos Verdadeiros (TP_i): Este termo refere-se ao componente M_{ii} , que representa a quantidade de casos corretamente identificados na categoria C_i . O cálculo dos itens na linha i , com exceção do M_{ii} , e representa o número total das diferentes classes que foram incorretamente classificadas como C_i . Falsos Negativos (FN_i): A soma dos elementos na linha i , exceto M_{ii} , que representa o número de instâncias da classe C_i que foram incorretamente classificadas como outra classe. Verdadeiros Negativos (TN_i): A soma de todos os elementos que não estão na linha i nem na coluna i , representando o número de instâncias que foram corretamente classificadas como não pertencentes à classe C_i .

A **Precisão** avalia a exatidão dos resultados positivos, ou seja, quantos dos casos classificados como positivos são, de fato, positivos. Essa métrica é especialmente importante quando os *falsos positivos* (classificar algo como positivo quando é negativo) têm um custo elevado, como em diagnósticos médicos ou filtragem de fraudes. A **Precisão** é calculada pela fórmula:

$$\text{Precisão}_i = \frac{TP_i}{TP_i + FP_i}$$

onde TP (Verdadeiros Positivos) representa o número de casos corretamente classificados como positivos, e FP (Falsos Positivos) representa o número de casos incorretamente classificados como positivos. Quanto maior a precisão, mais confiáveis são os resultados positivos fornecidos pelo modelo.

O **Recall** (ou *Sensibilidade*) mede a capacidade do modelo de identificar todas as instâncias positivas relevantes, sendo crucial em cenários onde *Falsos Negativos* são inaceitáveis. O **Recall** é calculado pela fórmula:

$$\text{Recall}_i = \frac{TP_i}{TP_i + FN_i}$$

onde TP (Verdadeiros Positivos) representa o número de casos corretamente classificados como positivos, e FN (Falsos Negativos) é o número de casos positivos que foram incorretamente classificados como negativos. Quanto maior o recall, maior a capacidade do modelo de identificar as instâncias positivas relevantes.

O **F1-Score** harmoniza as métricas de **Precisão** e **Recall**, sendo particularmente útil em cenários com classes desbalanceadas. Ele é a média harmônica entre a Precisão e o Recall, proporcionando uma métrica única que equilibra ambas, especialmente quando há uma preocupação em evitar altos valores de *Falsos Positivos* e *Falsos Negativos*. O **F1-Score** é calculado pela seguinte fórmula:

$$\text{F1-Score}_i = \frac{2 \times \text{Precisão}_i \times \text{Recall}_i}{\text{Precisão}_i + \text{Recall}_i}$$

Essa métrica é particularmente valiosa quando o objetivo é obter o equilíbrio entre a capacidade do modelo de identificar casos positivos (Recall) e a confiança nas classificações positivas (Precisão). Quanto maior o valor do F1-Score, melhor o modelo está equilibrando essas duas métricas.

O **Kappa de Cohen** é uma métrica que ajusta a concordância entre as previsões do modelo e os valores reais, levando em consideração a concordância que poderia ocorrer apenas pelo acaso. Em problemas de classificação, especialmente quando as classes estão desbalanceadas, a acurácia por si só pode ser enganosa, pois um modelo pode parecer ideal apenas por prever a classe majoritária na maior parte dos casos. O **Kappa de Cohen** corrige esse efeito ao comparar a concordância real com a concordância esperada pelo acaso. Sua fórmula é:

$$\kappa = \frac{p_0 - p_e}{1 - p_e}$$

onde: - p_0 representa a **precisão observada**, ou seja, a proporção de previsões corretas feitas pelo modelo. - p_e representa a **precisão esperada**, que é a concordância esperada caso as previsões fossem atribuídas aleatoriamente, apenas com base na distribuição das classes.

O termo "acaso" nesse contexto refere-se à situação em que as previsões do modelo não possuem real poder discriminativo e apenas seguem a distribuição das classes no conjunto de dados. Se p_0 for próximo de p_e , significa que o modelo não está aprendendo padrões reais e sua concordância com os rótulos verdadeiros não é melhor do que uma atribuição aleatória. Já valores mais altos de κ indicam que o modelo está capturando informações significativas e classificando corretamente além do que seria esperado por sorte.

Valores de κ variam entre -1 e 1, onde: - $\kappa = 1$ indica concordância perfeita entre previsões e rótulos reais. - $\kappa = 0$ indica que a concordância não é melhor do que o acaso. - $\kappa < 0$ sugere uma discordância sistemática, ou seja, o modelo está errando mais do que acertando.

A **Acurácia Global** (*Overall Accuracy - OA*) é uma métrica que indica o sucesso geral do modelo em todas as classes, representando a proporção de previsões corretas em relação ao total de amostras avaliadas. Sua fórmula é dada por:

$$\text{OA} = \frac{\sum_{i=1}^C M_{i,i}}{\sum_{i=1}^C \sum_{j=1}^C M_{i,j}}$$

onde: - C é o número total de classes no problema de classificação. - $M_{i,i}$ representa os elementos da diagonal principal da matriz de confusão, ou seja, os casos corretamente classificados para cada classe i . - $M_{i,j}$ representa os elementos da matriz de confusão, incluindo tanto classificações corretas quanto incorretas.

A OA mede a taxa de acertos do modelo considerando todas as classes, sendo intuitiva e amplamente utilizada. No entanto, em problemas com classes desbalanceadas, a acurácia global pode ser uma métrica enganosa, pois um modelo pode obter uma alta OA apenas por classificar corretamente a classe majoritária, sem necessariamente apresentar o desempenho nas classes minoritárias. Para uma avaliação mais equilibrada, métricas adicionais como Precisão, Recall, F1-Score e Kappa de Cohen são recomendadas.

A **Acurácia Média** (*Average Accuracy - AA*) mede a média das acurácias individuais de cada classe, sendo especialmente útil para avaliar o desempenho do modelo em conjuntos de dados desbalanceados. Diferente da *Acurácia Global* (*Overall Accuracy - OA*), que pode ser influenciada por classes majoritárias, a **AA** busca garantir uma avaliação mais justa entre todas as classes. Sua fórmula é dada por:

$$AA = \frac{1}{C} \sum_{i=1}^C \frac{TP_i}{TP_i + FN_i}$$

onde: - C é o número total de classes no problema de classificação. - TP_i representa o número de Verdadeiros Positivos para a classe i . - FN_i representa o número de Falsos Negativos para a classe i .

A AA calcula a média das taxas de acerto individuais por classe, garantindo que o desempenho seja avaliado de maneira mais equitativa entre todas as classes, independentemente de sua frequência nos dados. Dessa forma, ela é uma métrica recomendada para cenários onde o desbalanceamento de classes pode comprometer a interpretação da Acurácia Global (OA).

Valores mais altos de AA indicam que o modelo está performando bem em todas as classes de maneira equilibrada, enquanto valores baixos podem sugerir que o modelo está favorecendo algumas classes em detrimento de outras.

Enfim, a **Perda de Teste** (*Test Loss*) é uma métrica fundamental na avaliação do desempenho de um modelo de aprendizado de máquina após o processo de treinamento. Ela quantifica o erro cometido pelo modelo ao comparar suas previsões com os rótulos verdadeiros no conjunto de teste. A perda é calculada com base em uma **função de perda** escolhida, dependendo do tipo de problema. No caso de classificação, uma função comum é a **Entropia Cruzada**, dada por:

$$\text{Perda} = -\frac{1}{N} \sum_{i=1}^N y_i \log(\hat{y}_i)$$

onde: - N é o número total de amostras no conjunto de teste. - y_i representa o rótulo verdadeiro da i -ésima amostra. - $\hat{y}_i$ é a previsão do modelo para a i -ésima amostra.

A função de perda mede o quão distante as previsões do modelo estão dos valores reais. Durante o treinamento, o objetivo é minimizar essa perda, garantindo que o modelo faça previsões mais precisas. Uma perda de teste baixa sugere que o modelo generaliza bem para dados não vistos, enquanto uma perda elevada pode indicar problemas como underfitting ou overfitting.

A Perda de Teste deve ser analisada em conjunto com outras métricas, como **Acurácia, Precisão, Recall e F1-Score**, para fornecer uma avaliação completa do desempenho do modelo.

Essas métricas, em conjunto, oferecem uma visão abrangente da eficácia e robustez do modelo, permitindo uma análise detalhada de seu desempenho em diferentes aspectos.

4.5 Considerações Finais

Este capítulo expôs os materiais e métodos utilizados no desenvolvimento e avaliação do modelo, abrangendo desde a preparação e divisão dos dados até as especificações da arquitetura e estratégias de validação, com destaque para as técnicas de pré-processamento e a abordagem híbrida das CNNs, essenciais para superar os desafios do processamento de imagens hiperespectrais. A implementação da validação cruzada estratificada em cinco divisões assegurou uma avaliação justa e robusta do modelo, enquanto as métricas de avaliação adotadas proporcionaram uma análise detalhada e abrangente do desempenho, validando a eficácia do modelo proposto em variados contextos de dados.

A aplicação dos métodos descritos reforça a confiança nos resultados obtidos e estabelece uma base para futuras pesquisas no campo de sensoriamento remoto hiperspectral. Espera-se que contribuam para o avanço das técnicas de classificação e mapeamento agrícola, com possíveis futuras melhorias na arquitetura do modelo, ajustes no pré-processamento e expansão para outras aplicações de sensoriamento remoto.

5 Resultados e Discussão

Neste capítulo, são analisados e discutidos os resultados obtidos por meio da aplicação do modelo desenvolvido para classificar imagens de sensoriamento remoto hiperespectral em áreas agrícolas. Na comparação do desempenho do modelo com técnicas tradicionais e outras abordagens atuais de aprendizado profundo, avaliamos sua efetividade utilizando as seguintes métricas: precisão, recall, F1-Score, kappa, acurácia global, acurácia média e perda nos testes. Os experimentos foram executados por meio do ambiente Google Colab Pro ¹, o que disponibiliza recursos computacionais avançados, fundamentais para a realização deste estudo. Foi utilizado o ambiente computacional com a seguinte configuração: L4 GPU (53 GB RAM, RAM da GPU 22.5 GB e 78,2 GB de disco).

5.1 Resultados do Método Proposto

O desempenho do modelo proposto é mostrado na Tabela 10, que mostra que o modelo obteve acurácia entre 99,90% e 100,00% em todas as bases de dados testadas. As métricas para OA e AA se aproximaram de 100%, demonstrando a capacidade do modelo de generalizar os resultados para diversas bases de dados. A estatística Kappa também confirma esse desempenho, mostrando uma concordância entre as classificações verdadeiras e as especificidades do modelo, indicando que o modelo não sofreu superajuste.

Tabela 10 – Resultados do Modelo Proposto nos Testes nas Diversas Bases de Dados

Base de Dados	OA (%)	AA (%)	Kappa (%)	Precisão (%)	Recall (%)	F1-Score (%)
Indian Pines	99,90	99,93	100,00	99,90	99,90	99,90
Salinas	100,00	100,00	100,00	100,00	100,00	100,00
Universidade de Pavia	100,00	100,00	100,00	100,00	100,00	100,00
Centro Espacial Kennedy (KSC)	100,00	100,00	100,00	100,00	100,00	100,00
Botswana	100,00	100,00	100,00	100,00	100,00	100,00
WHU-Hi-LongKou	99,97	99,92	100,00	99,97	99,97	99,97
WHU-Hi-HanChuan	99,99	99,99	100,00	99,99	99,99	99,99
WHU-Hi-HongHu	99,96	99,89	100,00	99,96	99,96	99,96

Estes resultados atestam a efetividade do modelo, ressaltando sua utilidade em uma variedade ampla de situações de sensoriamento remoto. A habilidade do modelo de manter 100% de precisão em diversas bases de dados indica que ele pode ser uma ferramenta útil para usos práticos que refletem análises precisas e confiáveis de dados hiperespectrais.

A Figura 13 apresenta a relação entre o número de componentes principais (PCA) e a acurácia de classificação para diferentes bases de dados: Salinas, Indian Pines, Botswana, UniPavia, KSC e WHU-Hi. Note-se que a acurácia alcança um valor estável a partir de três componentes principais, independentemente do conjunto de dados. Esse tipo de

¹ <https://colab.google/>

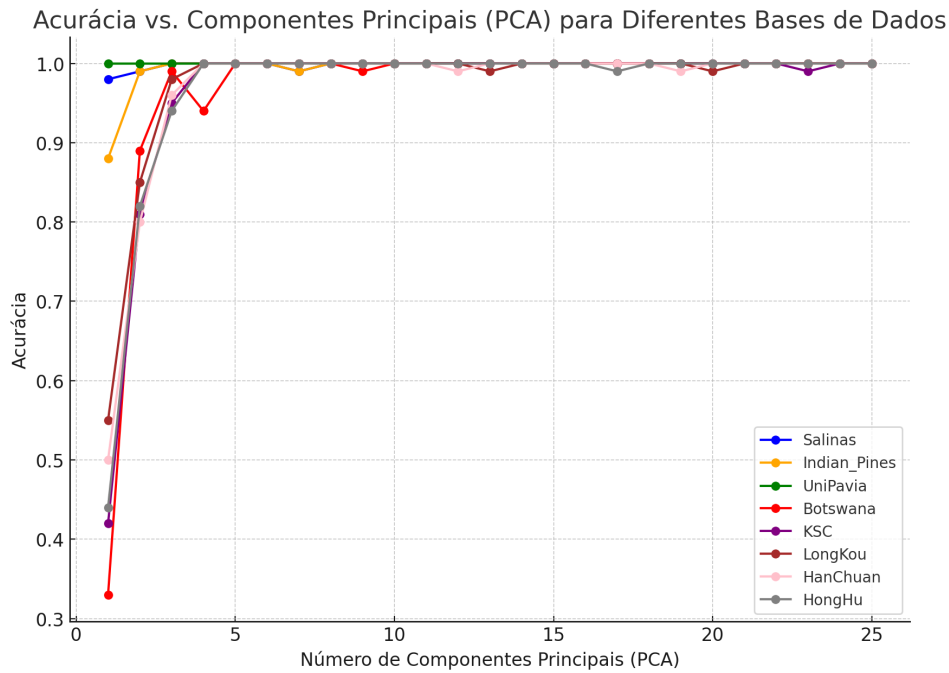


Figura 13 – Análise de Componentes Principais (PCA) para Redução de Dimensionalidade

comportamento sugere que três componentes principais são adequados para representar as características significativas dos dados, alcançando acurácia próxima a 100% em todas as bases examinadas. A partir deste momento, o acréscimo na quantidade de elementos não proporciona avanços no desempenho, o que sugere que os componentes adicionais incorporam informações redundantes ou irrelevantes para a atividade de classificação.

Assim, para o modelo proposto, a utilização de 3 componentes principais foi considerada ideal conforme a métrica analisada, pois oferece acurácia e simplicidade computacional. Essa escolha reduz a complexidade dos dados, diminui o consumo de memória e otimiza o tempo de processamento, mantendo, ao mesmo tempo, uma performance de classificação.

5.1.1 Indian Pines

Os resultados do modelo no treinamento na base de dados Indian Pines indicam um desempenho com uma acurácia global média nos k-folds de $100,00\% \pm 0,00$. Ademais, a perda registrada foi de $3,0000 \times 10^{-4}$, o que reflete um ajuste do modelo aos dados. O modelo no teste alcançou uma acurácia global de $99,90\%$ com uma perda de $2,8507 \times 10^{-3}$, reforçando sua confiabilidade e precisão.

Os gráficos na Figura 14 a evolução da perda e da acurácia do modelo ao longo das épocas durante o treinamento e a validação, utilizando a base de dados Indian Pines. No gráfico à esquerda, observa-se a Curva de Perda, onde a linha azul representa a perda durante o treinamento e a linha laranja mostra a perda na validação. Já no gráfico à direita, a Curva de Acurácia expõe que o modelo atingiu uma acurácia de 100% tanto

no treinamento quanto na validação logo no início, apontando que o modelo conseguiu classificar corretamente todas as amostras logo nas primeiras épocas, mantendo esse desempenho ao longo das demais. É importante ressaltar que o gráfico apresenta apenas 10 épocas devido à utilização do parâmetro de Early Stopping, que interrompeu o treinamento quando não foram observadas mais melhorias na métrica monitorada após um determinado número de épocas.

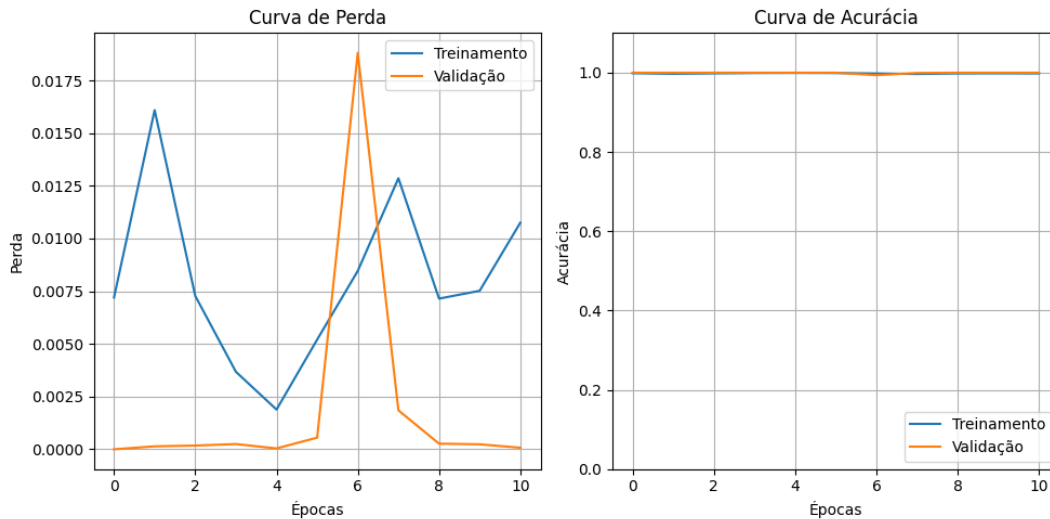


Figura 14 – Evolução da Perda e Acurácia do Modelo na Base Indian Pines

A Figura 15 ilustra uma comparação visual entre a verdade terrestre e a imagem prevista pelo modelo para a base de dados Indian Pines. À esquerda, a imagem mostra a distribuição real das classes de cobertura do solo, enquanto à direita temos as previsões do modelo. A correspondência próxima entre as duas imagens destaca a capacidade do modelo de reproduzir com precisão a distribuição espacial das diferentes classes. As áreas coloridas em ambos os mapas representam diferentes tipos de vegetação e uso do solo, com cada cor correspondendo a uma classe específica.

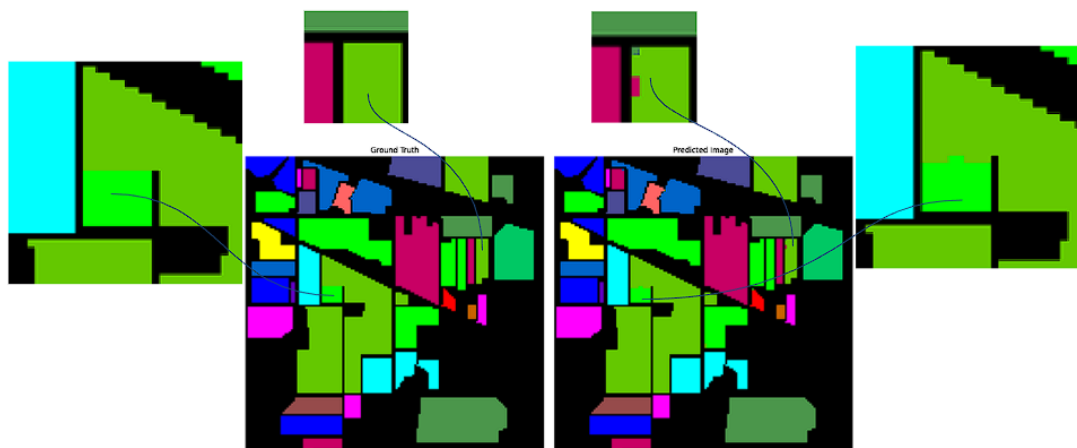


Figura 15 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo - Indian Pines

A análise visual dessas imagens permitiu avaliar qualitativamente o desempenho do modelo, proporcionando uma compreensão imediata das suas capacidades de classificação e de como ele maneja as variações dentro de uma cena hiperespectral. Podemos observar na imagem que a maioria das classes foi identificada corretamente, confirmando a viabilidade do modelo. Esse tipo de visualização é crucial para verificar a competência do modelo em termos práticos, indicando suas potencialidades e limitações.

5.1.2 Salinas

Os resultados do treinamento do modelo na base de dados Salinas mostram a acurácia global média de $100,00\% \pm 0.0$, apontando a robustez e a precisão do modelo. A perda foi de $1,4318 \times 10^{-6}$, destacando a aplicação do modelo em ajustar-se com precisão aos dados. Os resultados finais do teste do modelo na base de dados Salinas confirmam sua capacidade de classificação. O modelo alcançou uma perda de teste (Test Loss) de $1,0000 \times 10^{-7}$, o que indica uma ausência de erro na predição dos dados de teste.

A Figura 16 apresenta a evolução da perda e da acurácia do modelo durante o treinamento e a validação, utilizando a base de dados Salinas. No gráfico à esquerda, a Curva de Perda mostra a variação da perda ao longo das épocas. A linha azul representa a perda no treinamento, que apresenta uma oscilação por volta da quarta época, enquanto a linha laranja, correspondente à validação, se mantém consistentemente. Esses resultados sugerem que, embora o treinamento apresente algumas flutuações, o modelo conseguiu se ajustar aos dados de validação. No gráfico à direita, a Curva de Acurácia indica que o modelo atingiu uma acurácia de 100% tanto no conjunto de treinamento quanto no de validação desde o início do processo de treinamento, sustentando esse desempenho durante as 10 épocas. Assim como na base de dados anterior, este gráfico também mostra apenas 10 épocas devido à configuração de parada, que interrompe o treinamento quando não há mais melhorias na métrica de validação após um número pré-determinado de épocas.

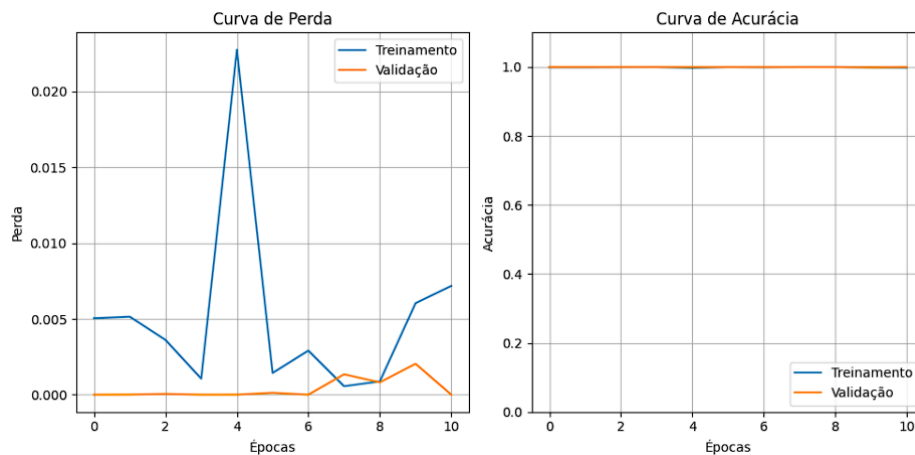


Figura 16 – Evolução da Perda e Acurácia do Modelo na Base Salinas

A Figura 17 apresenta uma análise visual comparativa entre a imagem de referência (*Ground Truth*) e a imagem resultante da predição pelo modelo (*Predicted Image*) para a base de dados Salinas. Esta comparação ilustrativa destaca a precisão do modelo na classificação das diferentes categorias de cobertura do solo, evidenciada pela correspondência exata entre as duas imagens. As áreas coloridas distintamente em cada painel representam diferentes classes de vegetação e uso do solo, onde cada cor corresponde a uma classe específica.

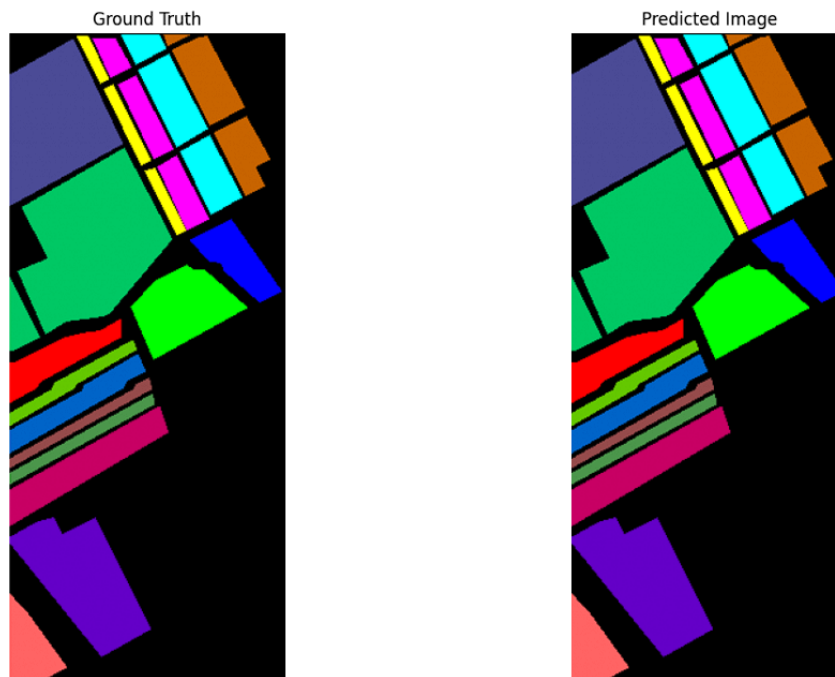


Figura 17 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo - Salinas

A acurácia do modelo é claramente demonstrada pelo alinhamento das divisões e cores entre a imagem de verdade terrestre e a imagem prevista. Tal precisão é indicativa do desempenho robusto do modelo na identificação correta das várias categorias, mesmo em regiões onde as classes de terreno são intercaladas. Esta capacidade de acertadamente classificar e mapear diferentes tipos de terreno é crucial para aplicações práticas em agricultura de precisão, gestão de recursos naturais e planejamento urbano.

5.1.3 Universidade de Pavia

A análise dos resultados do treinamento do modelo para a base de dados da Universidade de Pavia evidencia seu desempenho. A acurácia média alcançada após várias iterações de validação cruzada foi de $99,98\% \pm 0,05$. Esta consistência sugere que o modelo é estável e eficaz na classificação das classes. Além disso, a perda registrada foi de $8,0000 \times 10^{-4}$, demonstrando uma aptidão do modelo em minimizar o erro e alcançar uma convergência durante o treinamento.

Os resultados finais da aplicação do modelo no teste dos dados da Universidade de Pavia podem ser observados em seus indicadores de performance. A perda no teste foi de $8,0000 \times 10^{-7}$, ressaltando a precisão do modelo, com erros mínimos na identificação das classes durante a etapa de teste.. A acurácia global e a acurácia média atingiram ambos os 100% sugerindo que todas as classes foram classificadas corretamente sem qualquer exceção. Além disso, a estatística Kappa, que mede a concordância entre a predição e a verdade básica além do que seria esperado, também alcançou o valor máximo de 100% . Esses resultados enfatizam a valência do modelo em capturar e classificar corretamente as nuances dos dados hiperespectrais para aplicações em sensoriamento remoto.

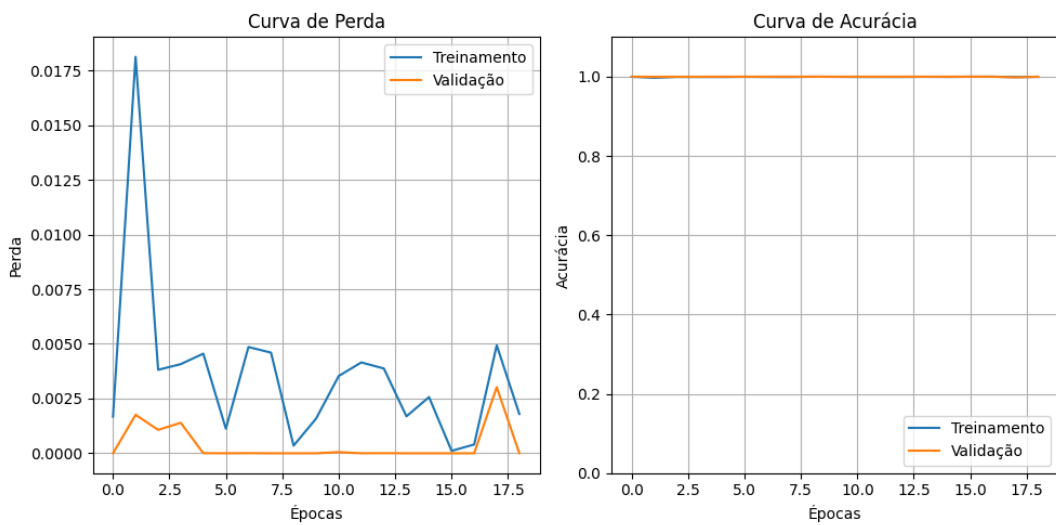


Figura 18 – Evolução da Perda e Acurácia do Modelo na Base Universidade de Pavia

A Figura 18 exibe a evolução da perda e da acurácia do modelo ao longo das épocas para a base de dados Universidade de Pavia. No gráfico à esquerda, a Curva de Perda ilustra a perda tanto no conjunto de treinamento quanto no de validação. A linha azul representa a perda no treinamento, que exibe algumas flutuações, especialmente na primeira época, antes de estabilizar seus valores. Já a linha laranja, correspondente à validação, mostra uma queda rápida, atingindo valores próximos de zero após a quarta época, apontando um ajuste eficaz do modelo nos dados de validação. No gráfico à direita, a Curva de Acurácia demonstra que o modelo atingiu 100% de acurácia tanto no conjunto de treinamento quanto no de validação desde as primeiras épocas, mantendo esse desempenho de forma estável até o final do treinamento. A Figura 18 mostra um total de 18 épocas, o que sugere que o modelo foi executado até que os critérios de Early Stopping fossem atingidos.

A imagem na Figura 19 faz uma avaliação visual comparativa entre a imagem de verdade terrestre (à esquerda) e a imagem prevista pelo modelo (à direita) para a base de dados da Universidade de Pavia. As cores vibrantes em ambas as imagens representam diferentes tipos de cobertura do solo, incluindo áreas urbanas, vegetação e corpos d'água,

com uma correspondência entre a classificação prevista e a realidade observada. Esta precisão visual robustece a praticabilidade do modelo na interpretação e classificação de dados hiperespectrais complexos, permitindo uma análise detalhada e acurada do uso do solo e da distribuição de diferentes materiais na área estudada. A imagem serve como um testemunho da capacidade do modelo em replicar com precisão a distribuição espacial das classes em um ambiente urbano e natural.

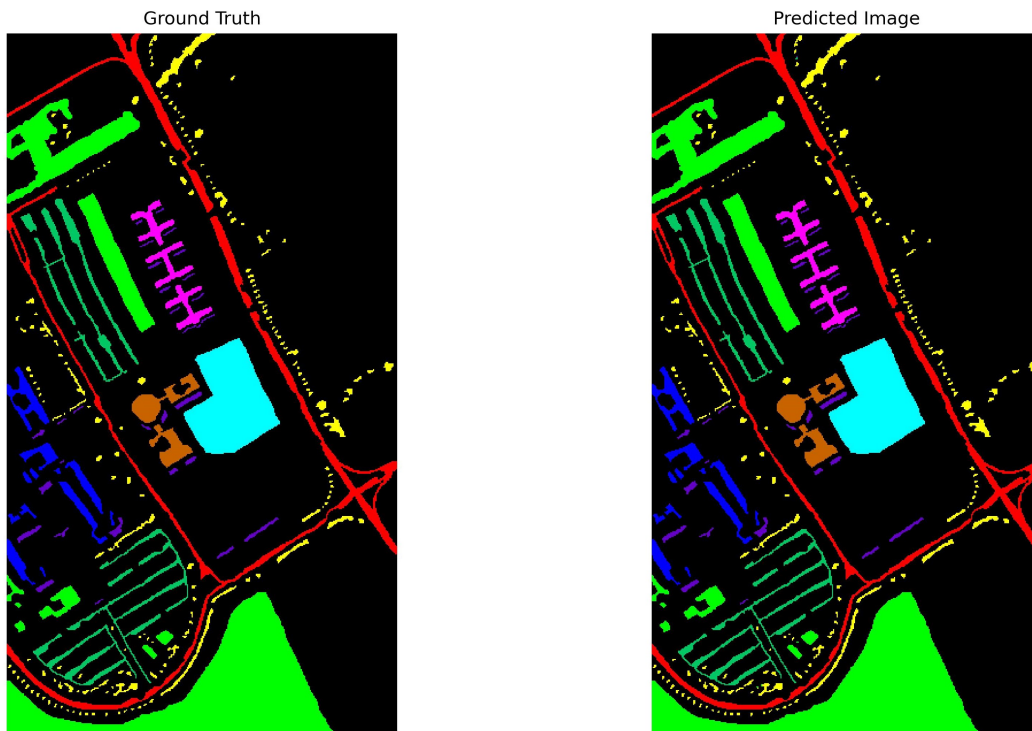


Figura 19 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo - Base Universidade de Pavia

5.1.4 Centro Espacial Kennedy (KSC)

Os dados de treinamento para a base Kennedy Space Center (KSC) mostram um desempenho geral do modelo, com uma acurácia global média de $100,00\% \pm 0.0$. O valor médio de perda durante o treinamento foi de $1,0000 \times 10^{-4}$, demonstrando uma convergência do modelo ao longo das iterações de aprendizado. Estes resultados ressaltam a utilidade do modelo em ajustar-se e generalizar sobre os dados desta base específica.

A eficácia do modelo é destacada no teste para a base Kennedy Space Center (KSC), apresentando uma eficiência do modelo, com todas as métricas alcançando o máximo possível. O modelo demonstrou perda de teste de $2,5500 \times 10^{-5}$, acurácia global, acurácia média, e estatística Kappa, todas atingindo $100,00\%$. Estes resultados refletem a precisão do modelo em classificar corretamente cada uma das categorias presentes nesta base de dados específica, reafirmando a solidez do modelo para aplicações práticas em sensoriamento remoto.

O gráfico na Figura 20 apresenta a evolução da perda e da acurácia do modelo ao longo das épocas para a base de dados KSC. No gráfico à esquerda, a Curva de Perda mostra as flutuações na perda tanto no conjunto de treinamento (linha azul) quanto no de validação (linha laranja). No treinamento, observa-se que a perda variou, atingindo picos em torno da sexta época, antes de começar a diminuir novamente. A linha laranja, que representa a validação, manteve-se quase constante e próxima de zero, o que sugere que o modelo conseguiu generalizar para os dados de validação. No gráfico à direita, a Curva de Acurácia revela que o modelo alcançou 100% de acurácia para ambos os conjuntos de dados logo nas primeiras épocas e manteve esse desempenho até a décima época. Este resultado indica que o modelo foi hábil na classificação das amostras da base KSC desde o início do treinamento. Assim como em outros experimentos, a quantidade de épocas foi limitada a 10, e não houve interrupção do treinamento antes desse ponto, como observado pelas flutuações no gráfico de perda durante todas as épocas.

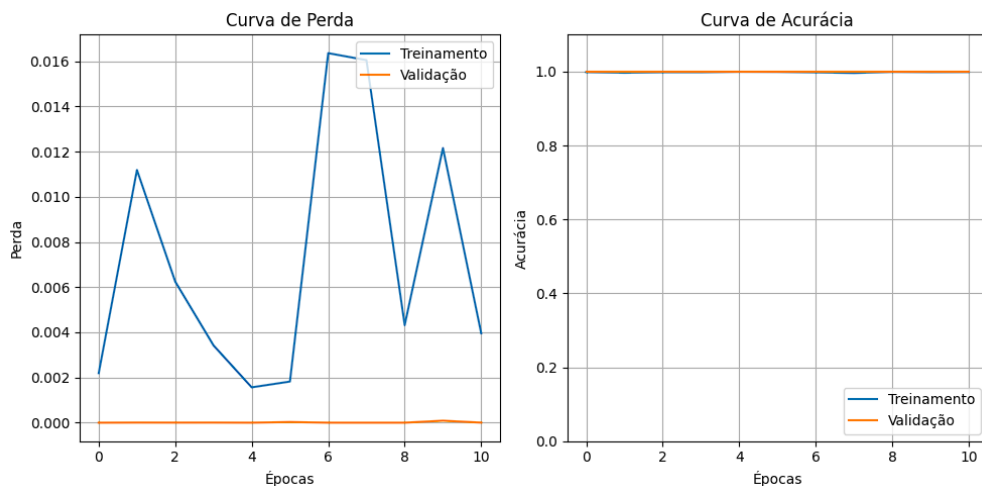


Figura 20 – Evolução da Perda e Acurácia do Modelo na Base KSC

As imagens na Figura 21 comparativas da base de dados Kennedy Space Center (KSC) demonstram a precisão do modelo na classificação de imagens hiperespectrais. À esquerda, a imagem *Ground Truth* mostra a distribuição real das classes vegetais e aquáticas da região, utilizando cores diversificadas para diferenciar claramente cada tipo de cobertura do solo. À direita, a imagem *Predicted Image* exhibe os resultados da classificação pelo modelo, onde é evidente a congruência com a verdade terrestre, destacando a praticabilidade do modelo em replicar com precisão os padrões de distribuição das classes na área observada. A semelhança visual entre as duas imagens reforça a acurácia do modelo na detecção e classificação dos diferentes tipos de vegetação e corpos d'água presentes no ecossistema do KSC.

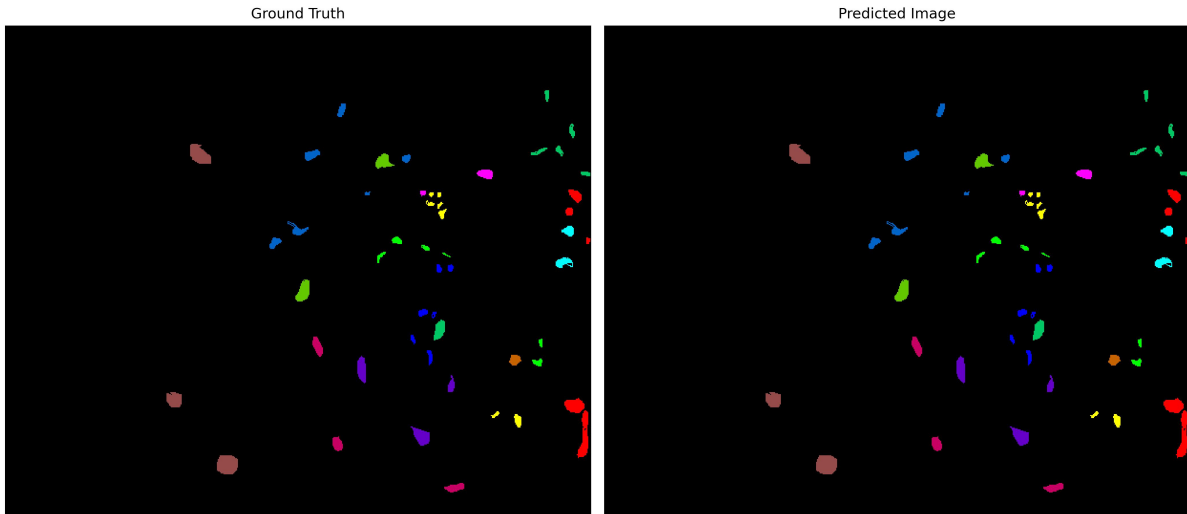


Figura 21 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo - Base KSC

5.1.5 Botswana

Os resultados do treinamento da base Botswana demonstram o desempenho do modelo com uma acurácia média de $99,96\% \pm 0.09$, refletindo uma estabilidade significativa durante as iterações de treinamento. A perda registrada foi de $1,0000 \times 10^{-3}$, indicando uma adaptação do modelo às características dos dados hiperespectrais. Estes resultados sublinham a firmeza do modelo em generalizar sobre os dados vistos durante o treinamento, como em dados novos avaliados durante a validação.

O desempenho do modelo no teste da base de dados Botswana apresenta uma perda de $4,3760 \times 10^{-4}$. A acurácia global e a acurácia média alcançaram 100% , destacando a precisão do modelo em prever corretamente todas as classes de terreno. Aliás, o coeficiente Kappa, que mede a concordância entre a classificação prevista e a verdadeira, também alcançou 100% , sugerindo uma correspondência total e uma capacidade do modelo em contexto real de teste.

Os gráficos da Figura 22 ilustram a evolução da perda e da acurácia do modelo durante o treinamento e validação, utilizando a base de dados Botswana. No gráfico à esquerda, a Curva de Perda mostra que a perda no conjunto de treinamento (linha azul) apresenta flutuações significativas, com picos observados na segunda e sexta épocas. Apesar dessas variações, a perda no conjunto de validação (linha laranja) se mantém praticamente estável e próxima de zero, com exceção da quarta época, mostrando que o modelo conseguiu ajustar-se aos dados de validação, minimizando erros. Já no gráfico à direita, a Curva de Acurácia expõe que o modelo alcançou uma acurácia de 100% em ambas as fases (treinamento e validação) logo nas primeiras épocas, mantendo-se constante até a décima

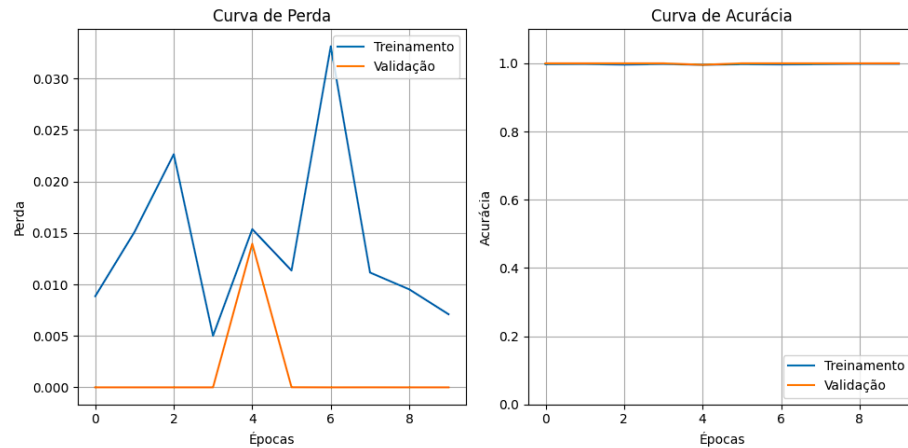


Figura 22 – Evolução da Perda e Acurácia do Modelo na Base Botswana

época. Isso sugere que o modelo teve um desempenho consistente, conseguindo classificar corretamente todas as amostras, tanto no conjunto de treinamento quanto no de validação.

A Figura 23 mostra duas representações visuais do mesmo conjunto de dados: à esquerda, a *Ground Truth* que representa a classificação correta das classes de vegetação e solo como mapeadas originalmente; à direita, a *Predicted Image* que exhibe o resultado da classificação realizada pelo modelo de aprendizado de máquina. Ambas as imagens usam uma paleta de cores específica para destacar as diferentes classes, o que facilita a comparação visual entre a precisão do modelo preditivo e a realidade do terreno. Esta comparação é essencial para validar a acurácia do modelo na tarefa de classificação de tipos de terreno e cobertura vegetal na área de estudo.

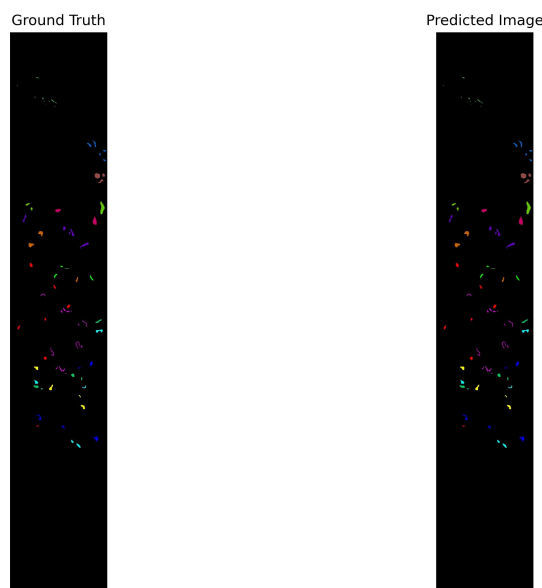


Figura 23 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo - Base de Dados Botswana

5.1.6 WHU-Hi-LongKou

Na base de dados LongKou, a acurácia média alcançada foi de $99,99\% \pm 0.02$, revelando uma consistência nos resultados através dos diferentes grupos de validação. O valor médio da perda foi de $7,0000 \times 10^{-4}$, refletindo a pretensão do modelo em classificar corretamente as amostras. Os resultados finais da predição na fase de teste para a base LongKou revelam a precisão do modelo. O modelo alcançou uma perda de teste de $4,3485 \times 10^{-3}$. A acurácia global também é de $99,97\%$, enquanto a acurácia média atingiu $99,92\%$. O coeficiente Kappa, que mede a concordância entre as previsões e as verdadeiras classificações, foi de $100,00\%$, demonstrando uma capacidade de generalização do modelo em dados de teste. Esses indicadores sugerem um desempenho robusto e confiável na classificação das imagens para a base de dados proposta.

Os gráficos apresentados na Figura 24 demonstram a progressão da perda e da acurácia do modelo ao longo do processo de treinamento e validação, utilizando a base de dados Long Kou. No gráfico à esquerda, a Curva de Perda apresenta um comportamento oscilante, tanto para o conjunto de treinamento (representado pela linha azul) quanto para o de validação (indicado pela linha laranja). Tal comportamento pode sugerir que o modelo está enfrentando dificuldades para estabilizar o aprendizado durante as fases de treinamento e validação. No gráfico à direita, a Curva de Acurácia indica que, embora houvesse variações na perda, o modelo foi capaz de atingir uma acurácia de 100% , tanto na fase de treinamento quanto na de validação. Isso indica que, apesar da instabilidade observada na perda, o modelo conseguiu classificar todas as amostras ao longo dos períodos específicos. Os gráficos indicam, igualmente, que o treinamento foi realizado por 12 épocas, e a acurácia permanece estável desde o princípio.

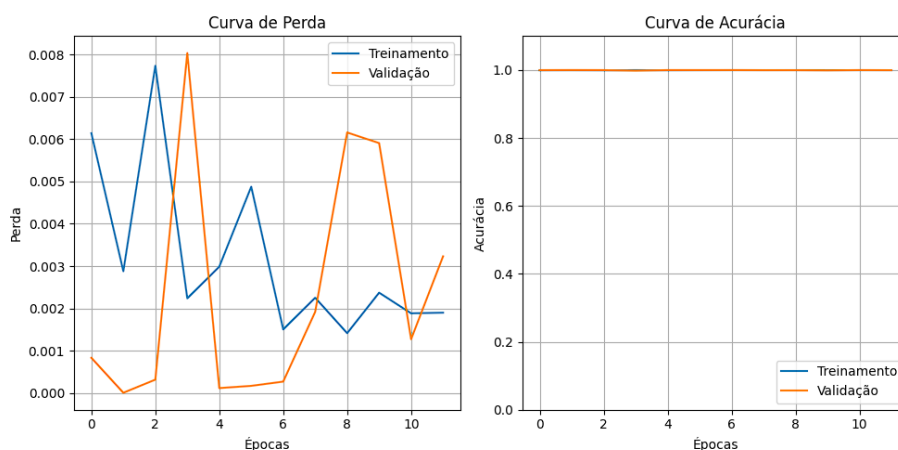


Figura 24 – Evolução da Perda e Acurácia do Modelo na Base LongKou

A imagem da Figura 25 apresenta os resultados de classificação de imagens de sensoriamento remoto da base de dados LongKou, demonstrando a verdade terrestre e a imagem predita lado a lado. As imagens revelam a precisão na classificação de diferentes culturas e terrenos, onde cada cor representa uma classe específica, como milho, algodão e

vários tipos de soja e terrenos expostos. Podemos afirmar que o modelo proposto generalizou o problema e gerou a classificação do mapa a partir de dados hiperespectrais. Esta visualização é crucial para verificar a aplicabilidade do modelo em replicar a distribuição espacial real das classes no terreno.

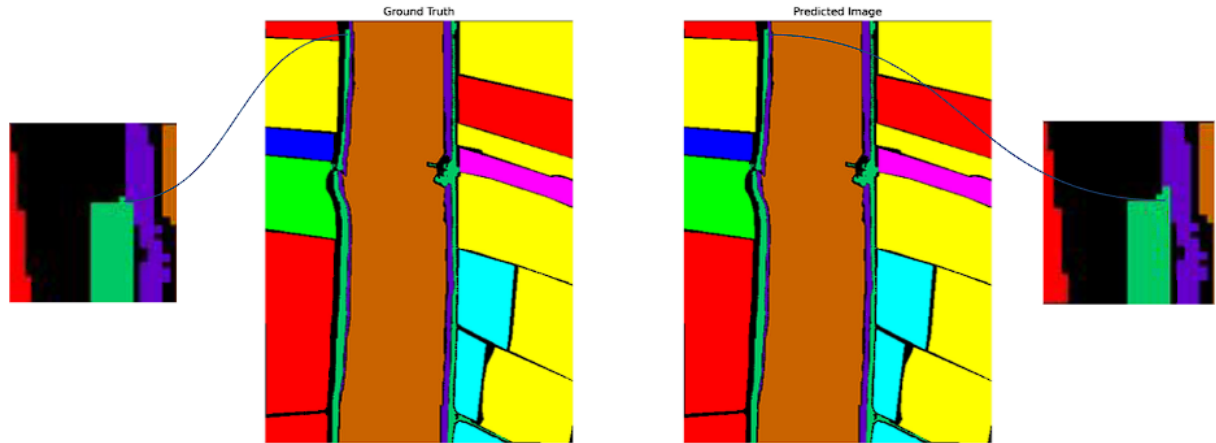


Figura 25 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo - Base LongKou

5.1.7 WHU-Hi-HanChuan

Os resultados de treinamento para a base de dados HanChuan destacaram uma valência do modelo, com uma acurácia de $99,99\% \pm (0.02)$, sinalizando uma consistência em diferentes grupos de validação. A perda de $4,0000 \times 10^{-4}$ indica que o modelo se ajustou de maneira competente aos dados, minimizando o sobreajuste e maximizando a capacidade de generalização. Este desempenho fortalece a solidez do modelo em condições de treinamento variadas, preparando-o para aplicações práticas eficientes e confiáveis.

Os resultados obtidos na fase de teste atestam a eficácia da abordagem adotada. A perda no teste foi de $6,1270 \times 10^{-4}$, refletindo uma precisão ideal na classificação dos dados. A acurácia global, Acurácia Média, atingiu $99,99\%$, assinalando que o modelo foi capaz de prever as classes com exatidão, demonstrando robustez e confiabilidade no processo de classificação das imagens hiperespectrais utilizadas.

Os gráficos da Figura 26 apresentam a evolução da perda e da acurácia do modelo durante o treinamento e validação na base de dados HanChuan. No gráfico à esquerda, vemos a Curva de Perda, onde a linha azul representa a perda durante o treinamento, enquanto a linha laranja corresponde à perda na validação. A perda do treinamento apresenta flutuações consideráveis, enquanto a perda da validação também apresenta variações, mas permanece geralmente abaixo da curva de treinamento, sugerindo ajuste do modelo, com a perda de validação sendo consistentemente menor. No gráfico à direita, a Curva de Acurácia mostra que o modelo atingiu $99,99\%$ de acurácia em ambas as fases

desde o início do processo de treinamento. A acurácia, tanto no treinamento quanto na validação, é constante ao longo de todas as 20 épocas, o que indica que o modelo é capaz de classificar corretamente as amostras no conjunto de validação e de treinamento sem apresentar dificuldades. Essas flutuações nas curvas de perda podem estar associadas à variação na complexidade das amostras processadas em cada época, mas a acurácia constante sugere que o modelo está aprendendo corretamente e generalizando para o conjunto de validação.

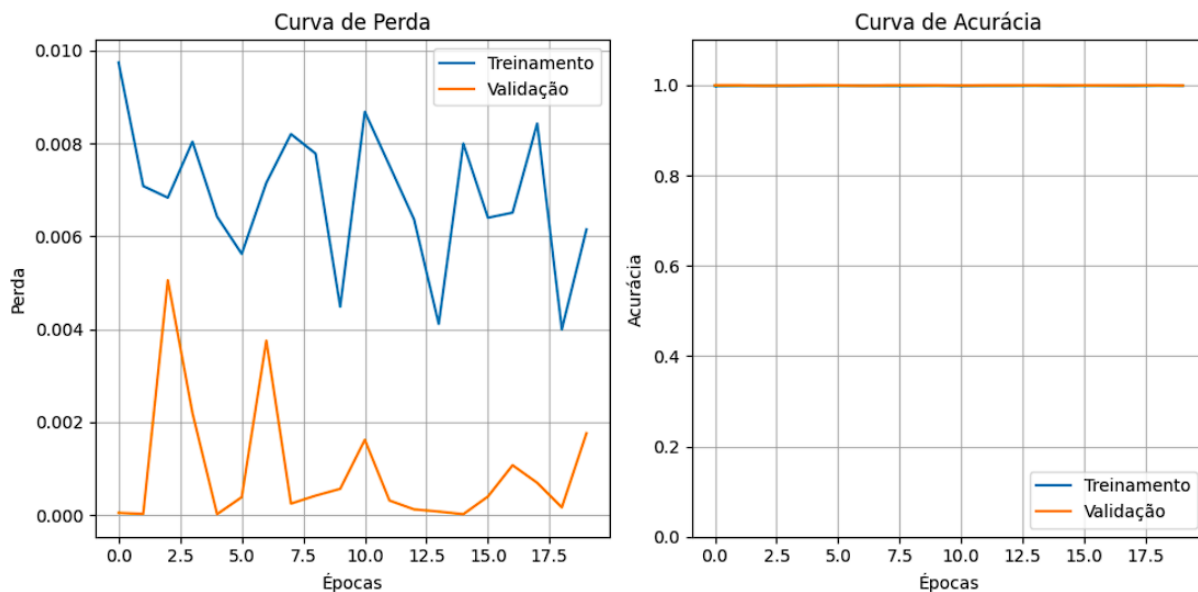


Figura 26 – Evolução da Perda e Acurácia do Modelo - Base HanChuan

As imagens apresentadas na Figura 27 demonstram o desempenho do modelo de classificação na tarefa de mapeamento de uso do solo e cobertura vegetal. Na imagem à esquerda, temos a verdade terrestre, que serve como referência para avaliar a precisão do modelo, enquanto a imagem à direita mostra os resultados preditos pelo modelo. A correlação entre as cores nas duas imagens indica que o modelo conseguiu identificar com precisão todas as classes presentes na área estudada, incluindo variações de vegetação e estruturas artificiais.

Os detalhes na imagem predita, especialmente em regiões onde diferentes tipos de culturas e elementos naturais se intercalam, ressaltam a firmeza do modelo em distinguir entre classes semelhantes, o que é um aspecto crucial para aplicações práticas em agricultura de precisão e planejamento urbano. Essa fidelidade nas previsões reforça a utilidade do modelo proposto para a análise de dados geospaciais em grande escala.

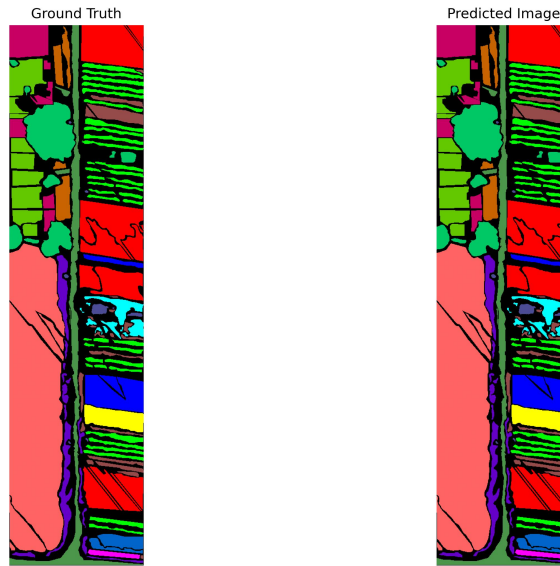


Figura 27 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo - Base HanChuan

5.1.8 WHU-Hi-HongHu

Os resultados de treinamento para a base de dados HongHu demonstram uma praticabilidade do modelo, alcançando uma acurácia média de aproximadamente $99,98\% \pm (0.01)$ assinalando consistência entre os diversos grupos. Para mais, a métrica de perda foi $6,0000 \times 10^{-4}$, o que sugere um ajuste efetivo do modelo aos dados sem incidência significativa de sobreajuste, assegurando a generalização do aprendizado.

O desempenho final do modelo no teste de predição para a base de dados na acurácia global foi de $99,96\%$ e média de $99,89\%$, com uma perda no teste de $1,5093 \times 10^{-3}$ e um coeficiente Kappa de $100,00\%$. Esses resultados indicam que o modelo conseguiu classificar quase todas as amostras de teste, apontando sua confiabilidade e eficácia.

Os gráficos da Figura 28 apresentam a evolução da perda e da acurácia do modelo ao longo das épocas na base de dados HongHu. No gráfico à esquerda, a Curva de Perda exibe as perdas durante o treinamento na linha azul e a validação na linha laranja. É possível observar que a perda de validação apresenta um pico na 15ª época, mas logo em seguida retorna para valores estabilizados, enquanto a perda no treinamento permanece relativamente constante ao longo do processo. No gráfico à direita, a Curva de Acurácia mostra que o modelo atingiu uma acurácia próxima de 100% tanto no conjunto de treinamento quanto no conjunto de validação, com uma pequena queda perto da 15ª época na validação. O modelo conseguiu manter a performance durante a maioria das épocas, o que sugere que ele foi capaz de generalizar para o conjunto de validação, apesar da pequena oscilação na 15ª época.

Esse pico na perda de validação pode ser decorrente de uma anomalia temporária

no conjunto de validação ou de amostras particularmente difíceis de classificar naquela época.

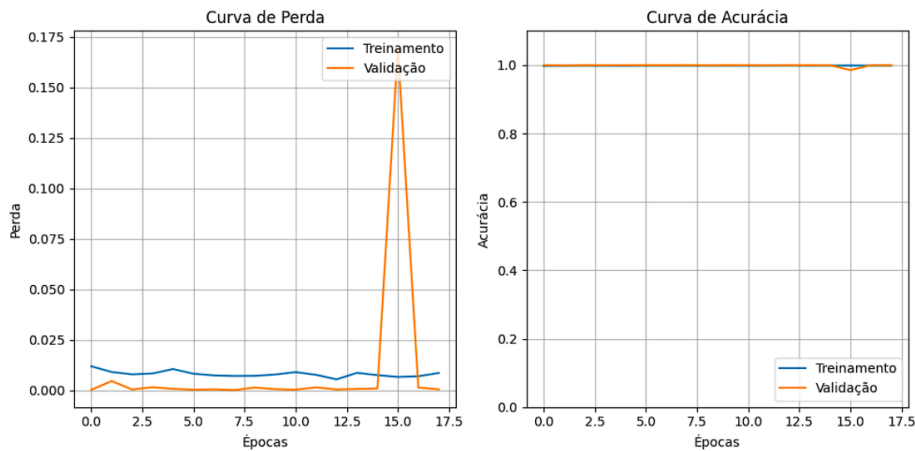


Figura 28 – Evolução da Perda e Acurácia do Modelo na Base HongHu

A imagem na Figura 29 ilustra uma comparação entre a *Ground Truth* e a *Predicted Image* para a classificação de múltiplas características de terreno na base de dados HongHu. À esquerda, a imagem verdadeira mostra a distribuição espacial real das diferentes classes das áreas agrícolas, diferentes tipos de vegetação e corpos d'água, coloridas vividamente para fácil distinção. À direita, a imagem predita pelo modelo exibe uma correspondência com a realidade, destacando a precisão do modelo em reproduzir as características observadas. A consistência das cores entre as duas imagens indica que o modelo conseguiu identificar e classificar corretamente as regiões, com cores correspondentes às da verdade terrestre, demonstrando sua aplicabilidade prática em tarefas de mapeamento e monitoramento via imagens de sensoriamento remoto.



Figura 29 – Comparação entre a Verdade Terrestre e a Imagem Prevista pelo Modelo - Base HongHu

5.2 Experimentos com o Modelo

Nesta seção, são descritos os experimentos executados com o modelo sugerido, com a finalidade de analisar seu desempenho em diversos cenários e conjuntos de dados. Os resultados obtidos por meio de testes com diferentes percentuais de amostras de treinamento são apresentados, juntamente com uma análise comparativa das principais considerações de avaliação. Além disso, procede-se a uma avaliação cruzada, na qual o modelo é treinado em um conjunto de dados e testado em outro, com o objetivo de verificar sua capacidade de generalização em diferentes distribuições de dados.

5.2.1 Experimento I - Desempenho com 10% de Amostras no Treinamento

Executar testes utilizando apenas 10% das amostras de treinamento é essencial para verificar a robustez do modelo em situações de dados restritos, uma condição frequente em aplicações práticas, como a classificação de imagens hiperespectrais, em que a entrega de grandes quantidades de dados rotulados é dispendiosa e prolongado. Essa metodologia possibilita a avaliação da habilidade de generalização do modelo ao operar com um conjunto limitado de dados, simulando cenários reais de limitações de amostras, principalmente em questões relacionadas a classes desequilibradas. Além disso, esta modalidade de teste permite verificar a eficácia das técnicas de regularização utilizadas, revelando se o modelo é capaz de manter um desempenho competitivo em condições desfavoráveis, aspecto que é fundamental para sua previsão prática em contextos de dados limitados.

A Tabela 11 expõe uma análise do desempenho dos modelos 2D-CNN, 3D-CNN utilizados no modelo de forma separadas e do Modelo Proposto juntas, empregando 10% das amostras para o treinamento, em quatro conjuntos de dados: Indian Pines, Salinas, Pavia Universidade e Centro Espacial Kennedy (KSC).

O modelo demonstrou um desempenho eficiente em todas as bases de dados, com exceção da base KSC, na qual o modelo 2D -CNN evidenciou-se ligeiramente eficaz nas métricas OA, Kappa e AA. Entretanto, o modelo sugerido se destacou nas bases Indian Pines e Salinas, obtendo uma acurácia global (OA) de 96,22% em Indian Pines e 99,95% em Salinas, superando os modelos 2D-CNN e 3D-CNN em todas as avaliações avaliadas.

Na base da Universidade de Pavia , o Modelo Proposto obteve resultados significativos, o 99,76% de acurácia global e 99,72% de acurácia média, sugerindo uma robusta habilidade de generalização em variados conjuntos de dados.

O desempenho limitado na base KSC pode estar associado a menor semelhança dos padrões espectrais dessa amostra em comparação com os outros, diminuindo que o modelo pode ser mais suscetível a dados com características específicas das que foram observadas durante o processo de treinamento.

Em resumo, o Modelo Proposto demonstrou uma capacidade robusta de aprendizado com uma quantidade limitada de dados de treinamento, destacando-se principalmente nas bases com maior volume de amostras, como Salinas e Pavia University.

Tabela 11 – Comparação de desempenho entre os Modelos 2D-CNN, 3D-CNN e o Modelo Proposto com 10% de amostras no treinamento

Método	Indian Pines			Salinas		
	OA (%)	Kappa (%)	AA (%)	OA (%)	Kappa (%)	AA (%)
2D-CNN	86,57	85,00	82,18	97,89	98,00	98,96
3D-CNN	86,02	84,00	74,11	99,19	99,00	99,68
Proposto	96,22	96,00	90,20	99,95	100,00	99,95
Método	Pavia University			Centro Espacial Kennedy (KSC)		
	OA (%)	Kappa (%)	AA (%)	OA (%)	Kappa (%)	AA (%)
2D-CNN	96,02	95,00	93,63	92,62	92,00	89,63
3D-CNN	98,73	98,00	98,85	82,03	80,00	73,61
Proposto	99,76	100,00	99,72	86,59	85,00	85,88

A Tabela 12 apresenta uma comparação de desempenho entre os modelos 2D-CNN, 3D-CNN e o Modelo Proposto, utilizando 10% das amostras de treinamento em quatro bases de dados: Botswana, WHU-Hi-LongKou, WHU-Hi-HanChuan e WHU-Hi-HongHu.

O modelo desenvolvido demonstrou um desempenho consistentemente eficiente, destacando-se nas métricas de OA, Kappa e AA. Em particular, nas bases Botswana e WHU-Hi-HanChuan, o modelo proposto alcançou 98,50% e 98,91% de OA, respectivamente, superando com ampla margem os modelos 2D-CNN e 3D-CNN. No caso da base WHU-Hi-LongKou, o modelo proposto também obteve a performance, com uma OA de 99,88% e um coeficiente Kappa de 100,00%, apontando uma concordância entre as previsões do modelo e os rótulos verdadeiros.

O desempenho inferior do modelo 3D-CNN em algumas bases, como Botswana, com 67,41% de OA, sugere que ele pode ser mais suscetível a dificuldades com a limitação de dados de treinamento, enquanto o modelo proposto se mostra mais robusto em condições adversas. Já o modelo 2D-CNN apresentou um desempenho intermediário, com uma performance aceitável, mas inferior ao modelo proposto em todas as bases de dados.

Em resumo, o Modelo Proposto demonstrou sua superioridade em termos de generalização, mesmo com um número reduzido de amostras de treinamento, revelando sua capacidade de adaptação a diferentes bases de dados e situações de dados limitados.

5.2.2 Experimento II - Análise Cruzada entre Conjuntos de Dados

Nesta subseção, analisamos a eficiência do modelo quando treinado em um conjunto de dados e avaliado em outro, com a finalidade de examinar sua habilidade de generalização e adaptação a distribuições específicas de dados. Tal avaliação é fundamental para

Tabela 12 – Comparação de desempenho entre os Modelos 2D-CNN, 3D-CNN e o Modelo Proposto com 10% de amostras no treinamento

Método	Botswana			WHU-Hi-LongKou		
	OA (%)	Kappa (%)	AA (%)	OA (%)	Kappa (%)	AA (%)
2D-CNN	91,69	91,00	91,76	99,36	99,00	98,37
3D-CNN	67,41	65,00	67,65	99,68	100,00	99,11
Proposto	98,50	98,00	98,67	99,88	100,00	99,70
Método	WHU-Hi-HanChuan			WHU-Hi-HongHu		
	OA (%)	Kappa (%)	AA (%)	OA (%)	Kappa (%)	AA (%)
2D-CNN	95,19	94,00	90,02	97,19	96,00	92,08
3D-CNN	96,72	96,00	94,35	97,65	97,00	93,93
Proposto	98,91	99,00	97,93	99,40	99,00	98,54

compreender o comportamento do modelo quando confrontado com situações que não foram apresentadas durante o processo de treinamento.

5.2.2.1 Avaliação do Modelo: Treinamento com Indian Pines e Teste com Salinas

Nesta fase, o modelo começou a ser aprimorado na base de dados *Indian Pines*, que abrange 16 categorias de tipos de vegetação e áreas de cultivo. Depois, o modelo foi avaliado na base *Salinas*, que também possui 16 classes, mas com características espectrais e espaciais distintas das observadas em *Indian Pines*. A finalidade desta análise foi verificar a habilidade do modelo de generalizar ao ser utilizado em um conjunto de dados distinto daquele onde foi treinado.

A Tabela 13 mostra os resultados obtidos durante o teste na base *Salinas*. Várias métricas foram utilizadas para avaliar o desempenho do modelo, incluindo a acurácia global, a acurácia média, estatística Kappa, precisão, recall e F1-Score.

Os achados sugerem que o modelo enfrentou desafios para generalizar corretamente os dados da base *Salinas*. Apenas 20,79% da acurácia total foi alcançada, com uma acurácia média de 6,26%. A estatística Kappa, que avalia a concordância entre as previsões do modelo e os rótulos reais, apresentou um valor de 11%, indicando uma correspondência reduzida, quase ao acaso. Ademais, a acurácia, o *recall* e o F1-Score são apresentados, revelando o desempenho geral do modelo na classificação das amostras de teste.

Esses resultados evidenciam que o modelo treinado em *Indian Pines* não foi capaz de capturar adequadamente as características espectrais e espaciais das classes presentes em *Salinas*, destacando a importância de técnicas de *fine-tuning* ou mesmo o treinamento direto em bases de dados mais próximas do domínio de teste.

Tabela 13 – Resultados do Modelo Treinado em Indian Pines e Testado em Salinas

Métrica	Valor
Overall Accuracy (OA)	20,79%
Average Accuracy (AA)	6,26%
Kappa Statistic	11,00%
Precision	12,21%
Recall	20,79%
F1 Score	15,21%

5.2.3 Experimento III - Avaliação Cruzada entre base de dados

Neste teste, o modelo foi inicialmente treinado utilizando a base de dados *Indian Pines*, um conjunto de dados hiperespectrais amplamente utilizado para tarefas de classificação de imagens. Esse treinamento permitiu ao modelo aprender as representações espectrais e espaciais associadas às diferentes classes agrícolas presentes nessa base de dados. A base *Indian Pines* consiste em imagens capturadas por sensores hiperespectrais, contendo 16 classes de diferentes tipos de cultivo e coberturas do solo.

5.2.3.1 Congelamento de Camadas

Após o treinamento inicial, foi realizado o **congelamento das primeiras 5 camadas do modelo**. Esse congelamento preserva o conhecimento adquirido nas camadas convolucionais iniciais, que capturam os padrões de baixo nível, como bordas e texturas, presentes nos dados da base *Indian Pines*. O congelamento das camadas evita que essas representações iniciais sejam alteradas durante o ajuste fino, permitindo que o modelo mantenha o aprendizado de características básicas enquanto se adapta a novos dados.

5.2.3.2 Fine-Tuning

Com as camadas iniciais congeladas, foi realizado o *fine-tuning* do modelo utilizando a base de dados *Salinas*. A base de dados *Salinas* também é composta por imagens hiperespectrais de alta resolução espacial e espectral, mas com classes agrícolas diferentes. Durante o *fine-tuning*, as camadas superiores do modelo (não congeladas) foram ajustadas para aprender os padrões específicos da nova base de dados, permitindo ao modelo adaptar-se aos novos tipos de cultivo presentes em *Salinas*. Esse processo de ajuste fino é uma técnica competente para aplicar o conhecimento previamente adquirido em uma base de dados a uma nova base com características semelhantes, melhorando o desempenho nas tarefas de classificação.

Tabela 14 – Resultados do Modelo Treinado em Indian Pines e Testado em Salinas com Ajuste Fino

Métrica	Valor (%)
Overall Accuracy (OA)	99,98
Average Accuracy (AA)	99,96
Kappa Statistic	100,00
Precision	99,98
Recall	99,98
F1 Score	99,98

5.3 Comparação com a Literatura

A performance do modelo proposto apresenta resultados promissores que são comparados com as abordagens tradicionais e de *deep learning* que são amplamente discutidas na literatura. Esta comparação é crucial para ilustrar a efetividade técnica do modelo. A seguir, são apresentados detalhadamente esses resultados, destacando as métricas onde o modelo demonstra melhorias significativas em relação aos métodos existentes.

A Tabela 15 apresenta um estudo de desempenho de várias abordagens de classificação aplicadas ao conjunto de dados Indian Pines. O método proposto destaca-se ao alcançar 99,90% em OA, 99,93% em AA, 100,00% em Kappa, e 99,90% em Precisão, Revocação e Medida F1. Os valores em negrito indicam os melhores resultados obtidos, sugerindo o êxito do método proposto na classificação da base de dados Indian Pines.

Tabela 15 – Tabela de comparação para a base Indian Pines

Método	OA (%)	AA (%)	Kappa(%)	Prec.(%)	Rec.(%)	F1 (%)
Agilandeewari et al. 2022	97,62	95,84	-	-	-	-
Vaishnavi et al. 2022	99,45	99,71	-	-	-	-
Zhong et al. 2017	99,19	98,93	99,07	-	-	-
Wang et al. 2022	98,81	98,83	98,65	-	-	-
Roy et al. 2019	99,75	99,63	99,71	-	-	-
Método Proposto	99,90	99,93	100,00	99,90	99,90	99,90

A Tabela 16 apresenta uma comparação de desempenho entre diferentes métodos de classificação aplicados à base de dados Salinas. O método proposto destaca-se ao alcançar 100% em todas as métricas aplicadas, visto que em Roy et al. 2019 são aplicadas 03 (três) métricas e 03 bases de dados. Sendo assim, o método proposto demonstra um desempenho confiável em relação aos métodos comparados, corroborando a efetividade do método proposto na classificação da base de dados Salinas.

A Tabela 17 expõe uma avaliação comparativa de métodos distintos de classificação utilizando a base de dados da Universidade de Pavia. O método proposto destaca-se ao

Tabela 16 – Tabela de comparação para a base Salinas

Métodos	OA (%)	AA (%)	Kappa(%)	Prec.(%)	Rec.(%)	F1 (%)
Patel et al. 2023	99,92	99,18	98,71	-	-	-
Agilandeewari et al. 2022	96,08	97,24	-	-	-	-
Vaishnavi et al. 2022	99,84	99,83	-	-	-	-
Zhong et al. 2020	90,05	94,86	88,92	-	-	-
Roy et al. 2019	100	100	100	-	-	-
Método Proposto	100	100	100	100	100	100

alcançar 100% em todas as métricas, demonstrando um desempenho eficaz em relação aos métodos comparados, indicando a validade do método proposto na classificação da base de dados da Universidade de Pavia.

Tabela 17 – Tabela de comparação para a base Universidade de Pavia

Métodos	OA (%)	AA (%)	Kappa(%)	Prec.(%)	Rec. (%)	F1 (%)
Patel et al. 2023	99,56	99,15	98,51	-	-	-
Vaishnavi et al. 2022	99,93	99,89	-	-	-	-
Zhong et al. 2017	99,79	99,66	99,72	-	-	-
Wang et al. 2022	99,76	99,60	99,69	-	-	-
Roy et al. 2019	99,98	99,97	99,98	-	-	-
Método Proposto	100	100	100	100	100	100

A Tabela 18 mostra uma análise comparativa de diferentes algoritmos de classificação na base de dados KSC. O método proposto destaca-se ao alcançar 100% em todas as métricas, apontando um desempenho capaz em relação aos métodos comparados, reforçando a vigência do método proposto na classificação da base de dados KSC.

Tabela 18 – Tabela de comparação para a base KSC

Métodos	OA(%)	AA(%)	Kappa(%)	Prec.(%)	Rec.(%)	F1(%)
Gao et al. 2018	92,22	-	95,60	-	-	-
Zhang et al. 2019	99,45	99,42	99,39	-	-	-
Nyabuga and Liu. 2021	99,48	99,48	99,42	-	-	-
Sofla and Shah-Hosseini. 2022	96,11	96,45	-	-	-	-
Zhong et al. 2017	99,61	99,33	99,56	-	-	-
Método Proposto	100	100	100	100	100	100

A Tabela 19 apresenta uma comparação de desempenho entre diferentes métodos de classificação aplicados à base de dados Botswana. O método proposto destaca-se ao alcançar 100% em todas as métricas, reforçando a efetividade do método proposto na classificação da base de dados Botswana.

Tabela 19 – Tabela de comparação para a base Botswana

Método	OA(%)	AA(%)	Kappa(%)	Prec.(%)	Rec.(%)	F1(%)
Gao et al. 2018	97,93	-	96,30	-	-	-
Nyabuga and Liu. 2021	99,87	99,89	99,86	-	-	-
Ge et al. 2020	96,44	96,80	96,14	-	-	-
Li et al. 2020	98,17	97,31	94,40	-	-	-
Sofla and Shah-Hosseini. 2022	100	100	-	-	-	-
Método Proposto	100	100	100	100	100	100

A Tabela 20 apresenta uma comparação de desempenho entre diferentes métodos de classificação aplicados à base de dados WHU-Hi-LongKou. O método proposto destaca-se ao alcançar os melhores resultados em todas as métricas avaliadas, com 99,97% em OA, 99,92% em AA, 100,00% em Kappa, e 99,97% em Precisão, Revocação e Medida F1. Os valores em negrito indicam os melhores resultados obtidos, demonstrando eficácia do método proposto em relação aos métodos comparados.

Tabela 20 – Tabela de comparação para a base WHU-Hi-LongKou

Métodos	OA (%)	AA (%)	Kappa(%)	Prec.(%)	Rec. (%)	F1 (%)
Zheng et al. 2020	99,17	-	-	-	-	-
Hamza et al. 2022a	99,23	-	96,08	96,54	96,53	-
Zhong et al. 2020	98,91	98,21	98,57	-	-	-
Zhang et al. 2023	99,65	99,06	99,54	-	-	-
Vaishnavi et al. 2022	99,91	99,60	-	-	-	-
Método Proposto	99,97	99,92	100,00	99,97	99,97	99,97

A Tabela 21 mostra uma análise de desempenho entre diversos métodos de classificação aplicados ao conjunto de dados WHU-Hi-HanChuan. O método proposto se destaca, alcançando 99,99% em quase todas as métricas, fortalecendo a confiabilidade do método proposto na classificação da base de dados WHU-Hi-HanChuan.

Tabela 21 – Tabela de comparação para a base WHU-Hi-HanChuan

Método	OA (%)	AA (%)	Kappa(%)	Prec.(%)	Rec. (%)	F1 (%)
Hamza et al. 2022a	97,15	-	75,64	77,25	77,09	-
Zheng et al. 2020	97,83	-	-	-	-	-
Zhong et al. 2020	93,95	92,69	92,90	-	-	-
Zhang et al. 2023	95,81	93,23	95,66	-	-	-
Vaishnavi et al. 2022	99,87	99,71	-	-	-	-
Método Proposto	99,99	99,99	100,00	99,99	99,99	99,99

A Tabela 22 exibe uma comparação da eficácia de diferentes técnicas de classificação na base de dados WHU-Hi-HongHu. O método proposto destaca-se ao alcançar entre 99,89%

e 100,00% em todas as métricas, indicando o êxito do método proposto na classificação da base de dados WHU-Hi-HongHu.

Tabela 22 – Tabela de comparação para a base WHU-Hi-HongHu

Método	OA (%)	AA (%)	Kappa(%)	Prec.(%)	Rec. (%)	F1 (%)
Xie et al. 2023	98,00	98,60	97,60	-	-	-
Zheng et al. 2020	97,45	-	-	-	-	-
Zhong et al. 2020	93,74	94,78	92,17	-	-	-
Zhang et al. 2023	97,40	93,13	96,71	-	-	-
Vaishnavi et al. 2022	99,06	98,12	-	-	-	-
Método Proposto	99,96	99,89	100	99,96	99,96	99,96

O modelo proposto supera os resultados das abordagens devido à sua capacidade de capturar informações contextuais, espaciais e espectrais de maneira mais eficaz. Ao integrar CNNs 3D com CNNs 2D, conseguimos explorar tanto as relações espaciais quanto as espectrais entre os pixels de uma imagem hiperespectral. Essa combinação permite uma classificação mais precisa, já que a CNN 3D pode aprender características complexas em múltiplas dimensões, enquanto a CNN 2D refina essas previsões aproveitando as informações extraídas pela CNN 3D.

5.4 Limitações

Embora os resultados encorajadores apresentados neste trabalho sejam notáveis, é essencial admitir algumas restrições que existem neste estudo, bem como o campo de classificação de imagens hiperespectrais de maneira geral. Os métodos elaborados estão subordinados de maneira substancial à qualidade e à quantidade dos dados hiperespectrais disponíveis. Além disso, sua habilidade de generalização para outros conjuntos de dados com características específicas ainda requer uma avaliação mais aprofundada.

A evolução do aprendizado profundo pode exigir recursos computacionais avançados, principalmente no que diz respeito ao processamento gráfico. Além disso, os modelos de aprendizado profundo funcionam frequentemente como caixas pretas, proporcionando escassa transparência quanto ao processo de tomada de decisões. Tal condição pode representar uma especificidade em contextos nos quais a interpretabilidade se mostra essencial. A eficácia dos modelos hiperespectrais geralmente está vinculada à seleção de hiper parâmetros, o que pode constituir uma atividade que demanda tempo e apresenta desafios técnicos, exigindo um entendimento abrangente tanto da teoria quanto da aplicação do modelo.

Os locais estudados por meio de sensoriamento remoto estão em constante transformação em razão de fatores naturais e atividades humanas. A compreensão dessas restrições pode direcionar investigações e ações futuras, transcender os desafios ainda presentes e

progressos no âmbito da classificação de imagens hiperespectrais. A habilidade de ajustar os modelos de classificação para que acompanhem essas alterações permanece sendo um desafio específico.

6 Conclusões e Trabalhos Futuros

Nesta pesquisa, investigou-se a integração de CNNs 3D e 2D para a classificação e o mapeamento de atividades agrícolas em imagens hiperespectrais. A abordagem proposta surgiu da necessidade de capturar tanto as características espectrais quanto espaciais dessas imagens de forma mais eficaz. Enquanto as CNNs 3D extraem informações espectrais contextualizadas, as CNNs 2D refinam os detalhes espaciais, resultando em um modelo mais robusto e preciso. Essa combinação demonstrou ser superior a abordagens exclusivamente 2D ou 3D, o que se refletiu nos resultados obtidos, com acurácia próxima de 100% em todas as bases avaliadas.

O desempenho do modelo pode ser atribuído à sua capacidade de explorar a redundância de informações nas imagens hiperespectrais, permitindo uma discriminação mais eficiente entre as classes agrícolas. Além disso, a abordagem combinada mitiga limitações comuns das técnicas individuais, garantindo uma representação mais fiel dos padrões espectro-espaciais.

Os resultados obtidos demonstram o cumprimento dos objetivos estabelecidos. A revisão da literatura possibilitou a identificação de lacunas e desafios na classificação de imagens hiperespectrais na agricultura, destacando a necessidade de metodologias mais eficazes. A aquisição e exploração de bases públicas viabilizaram a experimentação da abordagem proposta, enquanto o desenvolvimento de um método de pré-processamento melhorou a estruturação dos dados, otimizando a etapa de classificação. O modelo de CNN projetado integrou camadas 3D e 2D de forma eficiente, e sua validação, com ajustes na arquitetura e na distribuição dos dados, evidenciou sua capacidade de generalização em comparação com outras metodologias da literatura.

Dessa forma, este trabalho não apenas reforça a importância do uso de arquiteturas híbridas de aprendizado profundo na agricultura de precisão, mas também estabelece um ponto de partida para futuras pesquisas. Investigações adicionais podem explorar ajustes e otimizações para diferentes bases de dados, bem como expandir a aplicabilidade do modelo para outras áreas do sensoriamento remoto, consolidando seu impacto no monitoramento agrícola e ambiental.

Como trabalhos futuros, recomenda-se o aprimoramento do modelo visando melhorar sua capacidade de generalização, especialmente ao ser treinado em uma base de dados e testado em outra. Para isso, investigações adicionais podem explorar ajustes mais refinados dos hiperparâmetros e a adoção de técnicas avançadas de pré-processamento, como normalização aprimorada e redução de ruído mais eficaz. Além disso, testes com métodos alternativos de redução de dimensionalidade, como redes neurais autoencoders,

podem contribuir para uma representação mais eficiente das imagens hiperespectrais.

Outra direção promissora envolve o desenvolvimento de arquiteturas mais sofisticadas, como CNNs 3D/2D com mecanismos de atenção (*Transformers*) e modelos baseados em redes neurais generativas adversárias (*GANs*), que podem aumentar a capacidade do modelo de aprender representações mais discriminativas e adaptáveis. A avaliação do modelo em conjuntos de dados mais diversos, utilizando técnicas de aprendizado semi-supervisionado e aprendizado por reforço, pode ser fundamental para garantir uma melhor adaptação a novos cenários. Essas abordagens não apenas contribuirão para a melhoria da generalização do modelo, mas também para a otimização da complexidade computacional, interpretabilidade e robustez, possibilitando sua aplicação em condições mais desafiadoras e realistas.

Referências

AGILANDEESWARI, L. et al. Crop classification for agricultural applications in hyperspectral remote sensing images. *Applied Sciences*, MDPI, v. 12, n. 3, p. 1670, 2022. Citado 3 vezes nas páginas 8, 34 e 37.

BAABU, M. G.; PADMA, M. Semantic feature extraction method for hyperspectral crop classification. *Indonesian Journal of Electrical Engineering and Computer Science*, v. 23, n. 1, p. 387–395, 2021. Citado 2 vezes nas páginas 23 e 27.

BASCO, U. do P. *Hyperspectral Remote Sensing Scenes*. 2021. Disponível em: <https://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes#Pavia_Centre_and_University>. Citado 6 vezes nas páginas 12, 14, 15, 16, 18 e 40.

BHOSLE, K.; MUSANDE, V. Evaluation of deep learning cnn model for land use land cover classification and crop identification using hyperspectral remote sensing images. *Journal of the Indian Society of Remote Sensing*, Springer, v. 47, n. 11, p. 1949–1958, 2019. Citado 2 vezes nas páginas 34 e 37.

BHOSLE, K.; MUSANDE, V. Evaluation of cnn model by comparing with convolutional autoencoder and deep neural network for crop classification on hyperspectral imagery. *Geocarto International*, Taylor & Francis, v. 37, n. 3, p. 813–827, 2022. Citado 3 vezes nas páginas 8, 35 e 37.

COLKESEN, I.; OZTURK, M. Y. A comparative evaluation of state-of-the-art ensemble learning algorithms for land cover classification using worldview-2, sentinel-2 and rosis imagery. *Arabian Journal of Geosciences*, Springer, v. 15, n. 10, p. 942, 2022. Citado 2 vezes nas páginas 25 e 27.

DAOUD, E. A. Comparison between xgboost, lightgbm and catboost using a home credit dataset. *International Journal of Computer and Information Engineering*, v. 13, n. 1, p. 6–10, 2019. Citado na página 25.

DAVE, K.; VYAS, T.; TRIVEDI, Y. Band selection technique for crop classification using hyperspectral data. *Journal of the Indian Society of Remote Sensing*, Springer, v. 50, n. 8, p. 1487–1498, 2022. Citado 2 vezes nas páginas 25 e 27.

DEVLIN, J. et al. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018. Citado na página 29.

DIAKITE, A.; JIANGSHENG, G.; XIAPING, F. Hyperspectral image classification using 3d 2d cnn. *IET Image Processing*, Wiley Online Library, v. 15, n. 5, p. 1083–1092, 2021. Citado na página 12.

DUHAYYIM, M. A. et al. Automated deep learning driven crop classification on hyperspectral remote sensing images. *Computers, Materials & Continua*, v. 74, n. 2, 2023. Citado 2 vezes nas páginas 30 e 37.

FARMONOV, N. et al. Crop type classification by desis hyperspectral imagery and machine learning algorithms. *IEEE Journal of selected topics in applied earth observations*

and remote sensing, IEEE, v. 16, p. 1576–1588, 2023. Citado 3 vezes nas páginas 1, 31 e 37.

GALODHA, A. et al. Deep convolution neural networks with resnet architecture for spectral-spatial classification of drone borne and ground based high resolution hyperspectral imagery. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Copernicus GmbH, v. 43, p. 577–584, 2022. Citado 4 vezes nas páginas 7, 9, 35 e 37.

GALODHA, A. et al. Convolutional neural network (cnn) for crop-classification of drone acquired hyperspectral imagery. In: IEEE. *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing Symposium*. [S.l.], 2022. p. 7741–7744. Citado 4 vezes nas páginas 8, 9, 34 e 37.

GE, Z. et al. Hyperspectral image classification method based on 2d–3d cnn and multibranch feature fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, IEEE, v. 13, p. 5776–5788, 2020. Citado na página 1.

GEBBERS, R.; ADAMCHUK, V. I. Precision agriculture and food security. *Science*, American Association for the Advancement of Science, v. 327, n. 5967, p. 828–831, 2010. Citado na página 2.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. [S.l.]: MIT Press, 2016. Citado 2 vezes nas páginas 10 e 11.

HAMZA, M. A. et al. Squirrel search optimization with deep transfer learning-enabled crop classification model on hyperspectral remote sensing imagery. *Applied Sciences*, MDPI, v. 12, n. 11, p. 5650, 2022. Citado 2 vezes nas páginas 30 e 37.

HU, X. et al. S3anet: Spectral-spatial-scale attention network for end-to-end precise crop classification based on uav-borne h2 imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier, v. 183, p. 147–163, 2022. Citado 2 vezes nas páginas 29 e 37.

HU, Y. et al. Nondestructive classification of maize moldy seeds by hyperspectral imaging and optimal machine learning algorithms. *Sensors*, MDPI, v. 22, n. 16, p. 6064, 2022. Citado na página 8.

KARTHIKEYAN, L.; CHAWLA, I.; MISHRA, A. K. A review of remote sensing applications in agriculture for food security: Crop growth and yield, irrigation, and crop losses. *Journal of Hydrology*, Elsevier, v. 586, p. 124905, 2020. Citado na página 1.

KITCHENHAM, B. Procedures for performing systematic reviews. *Keele, UK, Keele University*, Citeseer, v. 33, n. 2004, p. 1–26, 2004. Citado na página 23.

LAFFERTY, J. et al. Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In: WILLIAMSTOWN, MA. *Icml*. [S.l.], 2001. v. 1, n. 2, p. 3. Citado na página 25.

LEI, L. et al. Docc: Deep one-class crop classification via positive and unlabeled learning for multi-modal satellite imagery. *International Journal of Applied Earth Observation and Geoinformation*, Elsevier, v. 105, p. 102598, 2021. Citado 2 vezes nas páginas 31 e 37.

- LIANG, L. et al. Multi-scale spectral-spatial attention network for hyperspectral image classification combining 2d octave and 3d convolutional neural networks. *Remote Sensing*, MDPI, v. 15, n. 7, p. 1758, 2023. Citado na página 29.
- LIANG, W. et al. Predicting hard rock pillar stability using gbd, xgboost, and lightgbm algorithms. *Mathematics*, MDPI, v. 8, n. 5, p. 765, 2020. Citado na página 25.
- LU, B. et al. Recent advances of hyperspectral imaging technology and applications in agriculture. *Remote Sensing*, MDPI, v. 12, n. 16, p. 2659, 2020. Citado na página 5.
- NANDIBEWOR, A.; HEGADI, R. A novel smlr-pso model to estimate the chlorophyll content in the crops using hyperspectral satellite images. *Cluster Computing*, Springer, v. 22, p. 443–450, 2019. Citado 2 vezes nas páginas 24 e 27.
- NIGAM, R. et al. Crop type discrimination and health assessment using hyperspectral imaging. *Current Science*, JSTOR, v. 116, n. 7, p. 1108–1123, 2019. Citado 3 vezes nas páginas 6, 24 e 27.
- NIU, B. et al. Hsi-transunet: A transformer based semantic segmentation model for crop mapping from uav hyperspectral imagery. *Computers and Electronics in Agriculture*, Elsevier, v. 201, p. 107297, 2022. Citado 2 vezes nas páginas 33 e 37.
- PATEL, U.; PATEL, S.; KATHIRIA, P. Hyperspectral image classification using transfer learning. In: *Communication and Intelligent Systems: Proceedings of ICCIS 2021*. [S.l.]: Springer, 2022. p. 545–556. Citado 2 vezes nas páginas 1 e 8.
- PATEL, U. et al. Crop type classification with hyperspectral images using deep learning: a transfer learning approach. *Modeling Earth Systems and Environment*, Springer, v. 9, n. 2, p. 1977–1987, 2023. Citado 4 vezes nas páginas 9, 33, 37 e 40.
- RICHARDS, J. A.; RICHARDS, J. A. et al. *Remote sensing digital image analysis*. [S.l.]: Springer, 2022. v. 5. Citado 2 vezes nas páginas 1 e 5.
- ROY, S. K. et al. Hybridsn: Exploring 3-d-2-d cnn feature hierarchy for hyperspectral image classification. *IEEE Geoscience and Remote Sensing Letters*, IEEE, v. 17, n. 2, p. 277–281, 2019. Citado 2 vezes nas páginas 1 e 11.
- SHAHROUDNEJAD, A. et al. Improved explainability of capsule networks: Relevance path by agreement. In: IEEE. *2018 IEEE Global Conference on Signal and Information Processing (globalsip)*. [S.l.], 2018. p. 549–553. Citado na página 31.
- SHINDE, P. P.; SHAH, S. A review of machine learning and deep learning applications. In: IEEE. *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*. [S.l.], 2018. p. 1–6. Citado na página 7.
- SINGH, P. et al. Crop type discrimination using geo-stat endmember extraction and machine learning algorithms. *Advances in Space Research*, Elsevier, 2022. Citado 2 vezes nas páginas 35 e 37.
- SU, W.-H. Advanced machine learning in point spectroscopy, rgb-and hyperspectral-imaging for automatic discriminations of crops and weeds: A review. *Smart Cities*, MDPI, v. 3, n. 3, p. 767–792, 2020. Citado na página 5.

- SUDHA, R.; VAIDEGHY, A. et al. Efficient weed classification and disease management using multivariate principle component analysis on spectral evolution using change detection methods. In: IEEE. *2022 IEEE International Conference on Data Science and Information System (ICDSIS)*. [S.l.], 2022. p. 1–6. Citado na página 1.
- THENKABAIL, P. S.; LYON, J. G.; HUETE, A. Advances in hyperspectral remote sensing of vegetation and agricultural crops. In: *Fundamentals, Sensor Systems, Spectral Libraries, and Data Mining for Vegetation*. [S.l.]: CRC press, 2018. p. 3–37. Citado 2 vezes nas páginas 2 e 5.
- VAISHNAVI, B. B. S. et al. Hyperspectral image classification for agricultural applications. In: IEEE. *2022 International Conference on Electronics and Renewable Systems (ICEARS)*. [S.l.], 2022. p. 1–7. Citado 5 vezes nas páginas 7, 9, 33, 37 e 40.
- WANG, A. et al. A hyperspectral image classification method based on adaptive spectral spatial kernel combined with improved vision transformer. *Remote Sensing*, MDPI, v. 14, n. 15, p. 3705, 2022. Citado 2 vezes nas páginas 32 e 37.
- WEI, L. et al. Precise crop classification using spectral-spatial-location fusion based on conditional random fields for uav-borne hyperspectral remote sensing imagery. *Remote Sensing*, MDPI, v. 11, n. 17, p. 2011, 2019. Citado 2 vezes nas páginas 26 e 27.
- WEI, L. et al. Spatial-spectral fusion based on conditional random fields for the fine classification of crops in uav-borne hyperspectral remote sensing imagery. *Remote Sensing*, MDPI, v. 11, n. 7, p. 780, 2019. Citado 2 vezes nas páginas 26 e 27.
- WU, H. et al. Precise crop classification of hyperspectral images using multi-branch feature fusion and dilation-based mlp. *Remote Sensing*, MDPI, v. 14, n. 11, p. 2713, 2022. Citado 4 vezes nas páginas 7, 8, 32 e 37.
- WU, Y. et al. Recognition of spider mite infestations in jujube trees based on spectral-spatial clustering of hyperspectral images from uavs. *Frontiers in Plant Science*, Frontiers Media SA, v. 14, p. 1078676, 2023. Citado 2 vezes nas páginas 24 e 27.
- XIE, J. et al. Hypersformer: A transformer-based end-to-end hyperspectral image classification method for crop classification. *Remote Sensing*, MDPI, v. 15, n. 14, p. 3491, 2023. Citado 2 vezes nas páginas 29 e 37.
- YANG, B.; HU, S. What kind of spatial and spectral resolution of uav-borne hyperspectral image is required for precise crop classification when using deep learning. In: IEEE. *2022 12th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing (WHISPERS)*. [S.l.], 2022. p. 1–5. Citado 4 vezes nas páginas 6, 9, 31 e 37.
- YANG, L. et al. Precise crop classification of uav hyperspectral imagery using kernel tensor slice sparse coding based classifier. *Neurocomputing*, Elsevier, v. 551, p. 126487, 2023. Citado 2 vezes nas páginas 26 e 27.
- YANG, L. et al. Kernel tensor sparse coding model for precise crop classification of uav hyperspectral image. *IEEE Geoscience and Remote Sensing Letters*, IEEE, v. 20, p. 1–5, 2023. Citado 3 vezes nas páginas 26, 27 e 28.
- YOU, L.; SUN, Z. Mapping global cropping system: Challenges, opportunities, and future perspectives. *Crop and Environment*, Elsevier, v. 1, n. 1, p. 68–73, 2022. Citado na página 1.

- YU, H. et al. Locality sensitive discriminant analysis for group sparse representation-based hyperspectral imagery classification. *IEEE Geoscience and Remote Sensing Letters*, IEEE, v. 14, n. 8, p. 1358–1362, 2017. Citado na página 29.
- YU, T. et al. S2-mlp: Spatial-shift mlp architecture for vision. In: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*. [S.l.: s.n.], 2022. p. 297–306. Citado na página 32.
- ZHANG, B. et al. Sanet: A self-attention network for agricultural hyperspectral image classification. *IEEE Transactions on Geoscience and Remote Sensing*, IEEE, 2023. Citado 2 vezes nas páginas 28 e 37.
- ZHANG, H. et al. Hyperspectral image classification on large-scale agricultural crops: The heilongjiang benchmark dataset, validation procedure, and baseline results. *Remote Sensing*, MDPI, v. 16, n. 3, p. 478, 2024. Citado 2 vezes nas páginas 30 e 37.
- ZHANG, H. et al. Detection of wheat fusarium head blight using uav-based spectral and image feature fusion. *Frontiers in plant science*, Frontiers Media SA, v. 13, p. 1004427, 2022. Citado na página 7.
- ZHANG, J. et al. Construction of a plant spectral library based on an optimised feature selection method. *Biosystems Engineering*, Elsevier, v. 195, p. 1–16, 2020. Citado 3 vezes nas páginas 7, 24 e 27.
- ZHANG, Y. et al. Topological structure and semantic information transfer network for cross-scene hyperspectral image classification. *IEEE Transactions on Neural Networks and Learning Systems*, IEEE, 2021. Citado na página 29.
- ZHAO, J. et al. A robust spectral-spatial approach to identifying heterogeneous crops using remote sensing imagery with high spectral and spatial resolutions. *Remote Sensing of Environment*, Elsevier, v. 239, p. 111605, 2020. Citado 3 vezes nas páginas 25, 27 e 28.
- ZHAO, X. et al. Multi-temporal hyperspectral classification of grassland using transformer network. *Sensors*, MDPI, v. 23, n. 14, p. 6642, 2023. Citado 2 vezes nas páginas 29 e 37.
- ZHONG, Y. et al. Whu-hi: Uav-borne hyperspectral with high spatial resolution (h2) benchmark datasets and classifier for precise crop identification based on deep convolutional neural network with crf. *Remote Sensing of Environment*, Elsevier, v. 250, p. 112012, 2020. Citado 9 vezes nas páginas 9, 1, 6, 10, 20, 21, 22, 32 e 37.