



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Luiz Cortinhas Ferreira Neto

**Método usando Deep Learning para processamento de
alvos de uso e cobertura da terra em imagens Landsat:
Estudo de caso da mineração no Brasil**

TD 06/2025

Belém
2025

Luiz Cortinhas Ferreira Neto

**Método usando Deep Learning para processamento de alvos
de uso e cobertura da terra em imagens Landsat: Estudo de
caso da mineração no Brasil**

Tese de Doutorado apresentado no Programa
de Pós-graduação em Engenharia Elétrica do
Instituto de Tecnologia como requisito parcial
para obtenção do título de Doutor em
Computação Aplicada.

TD 06/2025

Orientador: Aldebaro Barreto da Rocha Klautau Júnior

PPGEE/ITEC/UFPA

Belém- Pará – Brasil

2025

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)

C829m Cortinhas Ferreira Neto, Luiz.
Método usando Deep Learning para processamento de alvos de
uso e cobertura da terra em imagens Landsat: Estudo de caso da
mineração no Brasil / Luiz Cortinhas Ferreira Neto. — 2025.
115 f. : il. color.

Orientador(a): Prof. Dr. Aldebaro Klautau
Tese (Doutorado) - Universidade Federal do Pará, Instituto de
Tecnologia, Programa de Pós-Graduação em Engenharia Elétrica,
Belém, 2025.

1. Sensoriamento Remoto. 2. Computação. 3. Deep
Learning. 4. Mineração. 5. Brasil. I. Título.

CDD 006.312



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**“MÉTODO USANDO DEEP LEARNING PARA PROCESSAMENTO DE
ALVOS DE USO E COBERTURA DA TERRA EM IMAGENS LANDSAT:
ESTUDO DE CASO DA MINERAÇÃO NO BRASIL”**

AUTOR: LUIZ CORTINHAS FERREIRA NETO

TESE DE DOUTORADO SUBMETIDA À BANCA EXAMINADORA APROVADA PELO COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, SENDO JULGADA ADEQUADA PARA A OBTENÇÃO DO GRAU DE DOUTOR EM ENGENHARIA ELÉTRICA NA ÁREA DE COMPUTAÇÃO APLICADA.

APROVADA EM: 28/03/2025

BANCA EXAMINADORA:

Prof. Dr. Aldebaro Barreto da Rocha Klautau Júnior
(Orientador – PPGEE/ITEC/UFPA)

Prof. Dr. Claudomiro de Souza de Sales Júnior
(Avaliador Interno - PPGEE/ITEC/UFPA)

Prof^a. Dr.^a Alessandra Rodrigues Gomes
(Avaliadora Externa - INPE)

Prof. Dr. Nandamudi Lankalapalli Vijaykumar
(Avaliador Externo - INPE)

Prof. Dr. Washington de Jesus Sant’Anna Da Franca Rocha
(Avaliador Externo - UEFS)

VISTO:

Prof. Dr. Diego Lisboa Cardoso
(Coordenador do PPGEE/ITEC/UFPA)

À memória daqueles que sofreram e dos que se veem assolados pelos males causados pela atividade mineradora ilegal na Amazônia, e em especial, dedico este trabalho a Deus, fonte de toda sabedoria e inspiração; aos meus pais Luiz e Cassandra, pilares da minha vida e exemplos de amor incondicional; à minha esposa Ellen, companheira de jornada, por seu amor, paciência e apoio constante; e ao meu filho Luiz Caetano que me deu a energia que faltava para a conclusão dessa tese. A todos vocês, meu eterno agradecimento e a promessa de continuar lutando por um futuro mais justo e sustentável.

Agradecimentos

Primeiramente, meu coração transborda gratidão a meus amados pais. Obrigado por serem a base sólida sobre a qual construí meus sonhos. Seu amor incondicional, apoio incansável e sacrifícios incontáveis foram os pilares que me sustentaram durante toda a jornada acadêmica. Cada palavra de incentivo, cada abraço apertado e cada lágrima enxugada em momentos de dificuldade foram essenciais para que eu chegasse até aqui. Devo a vocês não apenas este título, mas a própria essência de quem me tornei.

Agradeço a Deus, fonte de toda sabedoria e força, por me guiar e iluminar meu caminho. Sua presença foi constante em cada etapa desta jornada, concedendo-me a resiliência para superar os obstáculos e a fé para acreditar no meu potencial. Agradeço pelas bênçãos recebidas, e pela oportunidade de crescer e aprender através desta experiência. Também sou imensamente grato a cada membro da Solved em especial a: Bruno Haick, Cesar Guerreiro, Maria Luize, e Julia Cansado, que doaram seu tempo para a leitura dessa tese. Obrigado por formarem uma equipe excepcional, um verdadeiro time de apoio e aprendizado mútuo. Agradeço pelas trocas de conhecimento, pelas risadas compartilhadas, pela amizade construída e pelo ambiente colaborativo que tornou essa jornada mais leve e gratificante. Vocês foram fundamentais para o meu desenvolvimento profissional e pessoal, e levo cada um de vocês com carinho no meu coração.

Agradeço também aos professores Raian Vargas Maretto e Claudio Persello, que compartilharam muito de seu tempo e conhecimento, e a toda a equipe do ITC/UT, que me ajudaram a enriquecer a discussão e o alcance desta tese.

Por fim, agradeço ao meu orientador Aldebaro e meu co-orientador, extra PPGEE, Pedro Walfir. Obrigado pela dedicação, pela perseverança e pela coragem de enfrentar os desafios que surgiram ao longo do caminho nada tradicional desta pesquisa. Agradeço pela confiança e suporte momentos de dúvida, pela disciplina em manter o foco e pela busca em sempre obter excelência, vocês são fontes de inspiração.

*“Onde a arara voava, agora o vazio,
Motor roncando forte, é um frio arrepio.
O rio cristalino, lamaçal barrento,
Peixe contaminado, fim do sustento.
Árvores gigantes, ao chão derrubadas,
Raízes expostas, almas arrancadas.
Biodiversidade, riqueza perdida,
No altar do ouro, a floresta é vendida.
O canto do uirapuru já não ecoa mais,
Silêncio de morte, em tristes funerais.
Garimpo ilegal, é sentença, em verde imensidão,
Apaga a vida, na terra e no coração.”*

Luiz Cortinhas F. Neto

Resumo

Técnicas de aprendizagem profunda, especificamente redes neurais convolucionais, têm grande potencial na análise de imagens de sensoriamento remoto para a detecção de mudanças na superfície terrestre. Nesta tese, uma metodologia automática baseada em uma CNN em formato de U (U-Net) é proposta para a detecção e mapeamento de áreas de mineração em todo o Brasil, diferenciando entre mineração industrial e garimpo (aluvião). A mineração, uma atividade extrativista que remove o substrato para acessar camadas ricas em sedimentos minerais, possui significativo impacto ambiental e socioeconômico, especialmente na Amazônia. Embora economicamente relevante, ainda não existem métodos automáticos para mapear essa atividade de forma contínua por longos períodos, distinguindo seus diferentes tipos. A metodologia proposta foi aplicada em mosaicos anuais Landsat livres de nuvens, abrangendo um período de 37 anos (1985-2022). A acurácia da U-Net modificada foi validada espacialmente por especialistas em sensoriamento remoto, obtendo-se uma média de 99% de acurácia geral, 91% de acurácia do produtor e 91% de acurácia do consumidor. Os resultados demonstram um aumento expressivo da área ocupada pela mineração, que cresceu cerca de 10 vezes entre 1985 e 2022, totalizando 4.500 km². O garimpo foi o subtipo de mineração que mais cresceu percentualmente, passando de 218 km² para 2.627km². A U-Net modificada e proposta se mostrou 30% menor em quantidade de parâmetros treináveis em comparação com U-Net original. Conclui-se que a metodologia proposta é eficaz para o mapeamento da atividade mineradora, fornecendo dados precisos e atualizados para a gestão ambiental e o planejamento territorial.

Palavras-chave: Deep Learning, U-Net, Redes convolucionais, Amazônia, Mineração.

Abstract

Deep learning techniques, specifically convolutional neural networks, have shown great potential in the analysis of remote sensing imagery for detecting changes in the Earth's surface. In this thesis, an automatic methodology based on a U-shaped CNN (U-Net) is proposed for the detection and mapping of mining areas in Brazil, differentiating between industrial mining and artisanal mining (alluvial). Mining, an extractive activity that removes the substrate to access layers rich in mineral sediments, has significant environmental and socioeconomic impacts, especially in the Amazon. Although economically relevant, there are still no automatic methods to map this activity continuously over long periods, distinguishing its different types. The proposed methodology was applied to annual cloud-free Landsat mosaics, covering a period of 37 years (1985-2022). The accuracy of the U-Net was spatially validated by specialists, obtaining an average of 99% overall accuracy, 91% producer accuracy, and 91% user accuracy. The results demonstrate a significant increase in the area occupied by mining, which grew about 10 times between 1985 and 2022, totaling 4500 km². Artisanal mining was the subtype of mining that grew the most proportionally, increasing from 218 to 2627 km². The proposed modified U-Net showed a 30% reduction in the number of trainable parameters compared to the original U-Net. It is concluded that the proposed methodology is effective for mapping mining activity, providing accurate and up-to-date data for environmental management and territorial planning.

Keywords: Deep Learning, U-Net, Convolutional Neural Networks, Amazon, Mining

Lista de ilustrações

Figura 1 - Planta de Mineração Industrial - Mosaico de imagens Landsat, fonte: O autor (2025)	23
Figura 2 – Mineração do tipo garimpeira - Mosaico de imagens Landsat, fonte: O autor (2025)	24
Figura 3 - Os três principais domínios de informação do sensoriamento remoto. Os domínios do espectro, do tempo e do espaço. Fonte: O autor (2025).	28
Figura 4 - Representação da Plataforma SatVeg (EMBRAPA-CNPTIA). Variação temporal de séries NDVI-MODIS sobre pastagem cultivada e agricultura anual de dois ciclos. Fonte: (SILVA <i>et al.</i> , 2022).....	29
Figura 5 - Características espectrais e temporais de corpos d’água. Em A, variação do MNDWI, em três alvos distintos. Em B, percentual de ocorrência de água superficial ao longo de 34 anos. Ainda em B, a tonalidade azul escuro representa ~100% de ocorrência, azul. Fonte: (DINIZ, C. <i>et al.</i> , 2021).....	30
Figura 6 - Data de lançamento de cada um dos satélites da família Landsat. Fonte: USGS.	32
Figura 7 - Características espectrais e espaciais dos sensores da família Landsat. O Multispectral Scanner System (MSS) a bordo do Landsats 1–5 tinha quatro bandas. O Thematic Mapper (TM) a bordo do Landsats 4 e 5 tinha sete bandas. O Enhanced Thematic Mapper Plus (ETM +) do Landsat 7 tem 8 bandas e o Landsats 8 e 9 com 11 bandas. Fonte: USGS.....	33
Figura 8 - O AM pode ser categorizado em supervisionado e não-supervisionado. Quando o classificador estima padrões gerais usando de instâncias preexistentes (amostras), temos o aprendizado supervisionado. Quando não há um resultado comparativo ou rótulo pré-existente, anterior a extração de atributos, considera-se o aprendizado como não-supervisionado. Fonte: O autor (2025).	35

Figura 9 - O algoritmo K-means separa n pontos de dados em k clusters. De tal forma que, cada elemento pixel do dado de entrada seja classificado em um cluster que tenha a menor diferença entre o valor médio (centróide) do cluster e o valor do pixel em questão. Fonte: O autor (2025).	38
Figura 10 - Subdivisão hierárquica dos algoritmos baseados em redes convolucionais. Fonte: O autor (2025), traduzido e ajustado de (EMEK SOYLU <i>et al.</i> , 2023).	44
Figura 11 - Comparativo entre algoritmos de DL, treinamento, métrica de Loss por época. Fonte: O Autor (2024).	56
Figura 12 - Comparativo entre algoritmos de DL, validação, métrica de Loss por época. Fonte: O Autor (2024).	58
Figura 13 - Comparativo entre algoritmos de DL, treinamento, métrica de <i>Binary IoU</i> por época. Fonte: O Autor (2024).	59
Figura 14 - Comparativo entre algoritmos de DL, validação, métrica de <i>Binary IoU</i> por época. Fonte: O Autor (2024).	60
Figura 15 - Arquitetura original da U-Net. Fonte: (RONNEBERGER; FISCHER; BROX, 2015).	63
Figura 16 - Estrutura da “U-Net modificada” utilizada para classificação neste trabalho. Fonte: O Autor (2024).	64
Figura 17 - Procedimento para validação da arquitetura modificada. Fonte: O Autor (2024).	65
Figura 18 - Gráfico comparativo de variação percentual de área entre o predito anual da U-Net modificada e a U-Net original de (RONNEBERGER; FISCHER; BROX, 2015). Fonte: O Autor (2024).	66
Figura 19 - Metodologia Aplicada para o Mapeamento das Minerações no Brasil. Fonte: O Autor (2024).	67
Figura 20 - Distribuição espacial da mineração no Brasil, apresentado em 5 gráficos: A) Distribuição espacial em 1985; B) Distribuição espacial em 2022; C) Área total da mineração por ano e por tipo; D) Os 5 estados com maior área de mineração em 2022; E) Os 5 municípios com maior área de mineração de garimpo em 2022. Fonte: O Autor (2025).	73
Figura 21 - Garimpos em áreas restritas, 3 gráficos: A) Serie temporal da expansão do garimpo em áreas protegidas; B) Mapa com os garimpos em áreas protegidas no Brasil; C) 5	

Terras indígenas e 5 Unidades de Conservação com maior presença de garimpo em 2022; e D) Extensão do garimpo dentro e fora das áreas protegidas. Fonte: O Autor (2024).	76
Figura 22 - Estratégia de amostragem aleatória e duplamente estratificada por célula. Fonte: O Autor (2024).	79
Figura 23- Série Temporal de avaliação da acurácia. Fonte: O Autor (2024).	80
Figura 24 - Análise espacial da acurácia, em 2022, possui quatro partes: A) Mapa da Acurácia Geral; B) Mapa da Acurácia do Consumidor; C) Mapa da acurácia do produtor; e D) Matriz de confusão média dos grids. Fonte: O Autor (2024).	81
Figura 25 - Terra Indígena Kayapó e garimpo em Vermelho, maior garimpo de extensão em linha reta, são 20km. Fonte: O Autor (2024).	86

Lista de tabelas

Tabela 1 - Informações gerais sobre cada um dos satélites.....	32
Tabela 2 - Características espaciais e espectrais dos sensores a bordo dos Landsat 5, Landsat 7. Fonte: USGS.	34
Tabela 3 - Algoritmos Citados no artigo de revisão da literatura Emek Soylu et al. (2023)	51
Tabela 4 - Os cinco algoritmos mais citados na literatura e os números de parâmetros treináveis em milhões	55

Lista de abreviaturas e siglas

DL	<i>Deep Learning</i>
AM	Aprendizado de Máquina
ALB	Amazônia Legal Brasileira
BQA	<i>Band Quality Assessment</i>
DPI	Processamento Digital de Imagem
AM	Aprendizado de Máquina
CNN	<i>Convolutional Neural Network</i>
IDH	Índice de Desenvolvimento Humano
TI	Terra Indígena
UC	Unidade de Conservação
RNN	<i>Recurrent Neural Network</i>
ERTS-1	<i>Earth Resources Technology Satellite</i>
MSS	<i>Multispectral Scanner System</i>
TM	<i>Thematic Mapper</i>
ETM+	<i>Enhanced Thematic Mapper Plus</i>
SWIR	<i>Short-Wave Infrared</i>
NIR	Near-Infrared
TIRS	<i>Thermal Infrared Sensor</i>
DNN	<i>Deep Neural Networks</i>
SVM	<i>Support Vector Machine</i>
LSTM	<i>Long Short Term Memory Cell</i>
GAN	<i>Generative Adversarial Networks</i>

WSSS	<i>Weakly-Supervised Semantic Segmentation</i>
GEE	<i>Google Earth Engine</i>
PDI	Processamento Digital de Imagens
PIB	Produto Interno Bruto
ANM	Agência Nacional de Mineração
NDWI	<i>Normilized Difference Water Index</i>
NDVI	<i>Normalized Difference Vegetation Index</i>
NDSI	<i>Normilized Difference Soil index</i>
MODIS	<i>Moderate Resolution Imaging Spectroradiomete</i>
MNDWI	<i>Modified Normalized Difference Water Index</i>
USGS	<i>United States Geological Survey</i>
MLP	<i>Multi Layer Perceptron</i>
DNN	<i>Deep Neural Network</i>
SVM	<i>Support Vector Machine</i>
GPU	<i>Graphics Processing Unit</i>
FPGA	<i>Field-Programmable Gate Array</i>
LSTM	<i>Long Short-Term Memory Cell</i>
GAN	<i>Generative Adversarial Networks</i>
ViT	<i>Vision Transformer</i>
ResNet	<i>Residual Network</i>
VGG	<i>Visual Geometry Group</i>
FSSS	<i>Fully-Supervised Semantic Segmentation</i>
WSSS	<i>Weakly-Supervised Semantic Segmentation</i>
FCNN	<i>Fully Convolutional Neural Network</i>
FLOPs	<i>Floating Point Operations</i>

VRAM	<i>Video Random-Access Memory</i>
SIGMINE/	Sistema de Informação Geográfica da Mineração
DNPM	Departamento Nacional de Produção Mineral
GeoTIFF	<i>Georeferenced Tagged Image File Format</i>
RMSE	<i>Root Mean Square Error</i>
Binary IoU	<i>Binary Intersection over Union</i>
O.A	<i>Overall Accuracy</i>
A.P	<i>Producers Accuracy</i>
A.C	<i>Consumers Accuracy</i>
IBAMA	Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis
SEMA	Secretaria Estadual do Meio Ambiente
ASM	<i>Artisanal and Small-Scale Mining</i>
ESA	<i>European Space Agency</i>
CONAMA	Conselho Nacional do Meio Ambiente
APA	Area de Proteção Permanente
FLONA	Floresta Nacional
RESEX	Reserva Extrativista
RPPN	Reserva Particular do Patrimonio Natural

Sumário

1	INTRODUÇÃO	18
1.1	PROBLEMÁTICA E JUSTIFICATIVA	20
1.2	CARACTERIZAÇÃO DA CLASSE ALVO	22
1.3	OBJETIVO	24
1.4	ESTRUTURA DA TESE	25
2	FUNDAMENTAÇÃO TEÓRICA	26
2.1	MINERAÇÃO NO BRASIL	26
2.2	SENSORIAMENTO REMOTO E SEUS DOMÍNIOS	27
2.3	PROGRAMA LANDSAT	31
2.4	APRENDIZADO DE MÁQUINA	34
2.5	CLUSTERIZAÇÃO COM K-MEANS	36
2.6	<i>DEEP LEARNING</i>	39
2.7	<i>DEEP LEARNING</i> APLICADO AO SENSORIAMENTO REMOTO	40
2.8	SEGMENTAÇÃO SEMÂNTICA	41
3	ARQUITETURAS DE DEEP LEARNING PARA SEGMENTAÇÃO SEMÂNTICA: ESTADO DA ARTE E CRITÉRIOS DE ESCOLHA	43
3.1	ESTADO DA ARTE DE MODELOS DE SEGMENTAÇÃO SEMÂNTICA	43
3.2	SEGMENTAÇÃO SEMÂNTICA APLICADA AO SENSORIAMENTO REMOTO	47
3.3	SELEÇÃO DE MODELOS PARA COMPARAÇÃO COM O MODELO PROPOSTO	49
4	MÉTODO PROPOSTO	62

4.1	MODIFICAÇÕES NA U-NET	62
4.2	ARQUITETURA HÍBRIDA PARA CLASSIFICAÇÃO TEMPORAL E EM LARGA ESCALA DA MINERAÇÃO NO BRASIL	67
5	<u>RESULTADOS</u>	<u>72</u>
5.1	EVOLUÇÃO DA ÁREA DE MINERAÇÃO: MINERAÇÃO INDUSTRIAL E GARIMPO	72
5.2	INDÍCIOS DE ILEGALIDADE E PERSISTÊNCIA TEMPORAL DE SÍTIOS GARIMPEIROS	75
5.3	ANÁLISE DE ACURÁCIA DO MAPA	77
6	<u>CONCLUSÕES</u>	<u>83</u>
6.1	DESEMPENHO E VALIDAÇÃO DA METODOLOGIA PROPOSTA	83
6.2	IMPORTÂNCIA SÓCIO-POLÍTICA DO DADO GERADO	84
6.3	TRABALHOS FUTUROS	89
7	<u>REFERÊNCIAS</u>	<u>90</u>

1 Introdução

A classificação digital de mudanças na superfície terrestre por meio de imagens de satélite é uma tarefa que tem se beneficiado enormemente dos avanços em técnicas de Processamento Digital de Imagem (DIP) e, mais recentemente, de aprendizagem profunda ou *deep learning* - DL. Neste contexto, a aplicação de redes neurais convolucionais (CNNs) para a análise de imagens de sensoriamento remoto surge como uma ferramenta capaz de lidar com a complexidade, diversidade e volume de dados envolvidos no monitoramento de vastas áreas ao longo do tempo. Esta tese se insere nesse cenário, propondo uma metodologia para o mapeamento da atividade de mineração em todo Brasil em larga escala espaço-espectro-temporal, utilizando a plataforma de computação em nuvem Google Earth Engine (GEE) e *deep learning*. O objetivo é reconhecer a evolução da mineração, diferenciando entre mineração industrial e garimpo (mineração artesanal), no período de 1985 a 2022, a partir de uma série histórica de 37 anos de dados orbitais (aproximadamente 20TB de dados cobrindo 4,5 milhões de km²).

A execução de um classificador nos três domínios possíveis da informação (tempo, espaço e espectro) aplicados ao monitoramento da mineração apresenta desafios específicos. Primeiro, a necessidade de distinguir variações potencialmente sutis no padrão de extração mineral em um ou mais desses domínios. Segundo a construção de uma metodologia que reduza o esforço computacional associado à aquisição, pré-processamento e pós-processamento de imagens digitais, operando dentro de limites financeiros viáveis para permitir a submissão do extenso conjunto de dados a um classificador.

A mineração consiste na extração de minerais e outros materiais geológicos de valor econômico da crosta terrestre. Essa atividade extrativista possui um papel histórico e economicamente relevante na economia global, especialmente em países da América do Sul (MACHADO; SILVIA, 2001). A regulamentação da mineração tem passado por transformações ao longo do tempo. Atualmente, a mineração industrial é predominante na maioria dos países. No entanto, a mineração artesanal, também conhecida como garimpo, tem apresentado expansão significativa nos últimos anos (CORTINHAS FERREIRA NETO *et al.*, 2024; SOUZA *et al.*, 2020). No Brasil, a relevância econômica dessa atividade não é diferente,

sendo responsável por quase 5% do PIB brasileiro (LEÃO; RABELO, 2023). Considerando apenas o primeiro semestre de 2021, aproximadamente 67% do saldo da balança comercial brasileira está diretamente relacionado a atividade extração mineral (LEÃO; RABELO, 2023). Agregando-se o saldo comercial ao retorno tributário, o montante econômico da atividade supera 210 bilhões de reais, em apenas 6 meses de operação, cenário que reforça relevância econômica da mineração nacional (LEÃO; RABELO, 2023).

Em contrapartida, a extração mineral, especialmente a garimpeira, é destacadamente prejudicial à saúde humana (VELÁSQUEZ RAMÍREZ *et al.*, 2021), resultado direto do uso tradicional de mercúrio (Hg) no processo de amalgamação do ouro. Mas para além deste, o garimpo amazônico está quase sempre associado a alterações ambientais e sociais ao longo dos ecossistemas terrestres. A atividade garimpeira é tradicional vetor de degradação florestal, contaminação cursos hídricos e, por consequência, de suas cadeias tróficas (CRESPO-LOPEZ *et al.*, 2021). Não é incomum que regiões garimpeiras, e suas vizinhas, apresentem altos índices de desmatamentos, elevadas concentrações de metilmercúrio (MeHg) em seus cidadãos e fauna, baixos índices de desenvolvimento humano (IDH), além elevada descarga sedimentar em seus cursos fluviais (BERZAS NEVADO *et al.*, 2010; TELMER *et al.*, 2006).

É importante ressaltar, porém, que minerar não é um crime em si mesmo. A extração mineral é atividade econômica salvaguardada pela constituição federal de quase todos os países Pan-Amazônicos (CORTINHAS FERREIRA NETO *et al.*, 2024). O crime, em geral, está ligado a posição no espaço quando é associado a invasão de terras protegidas (Terras Indígenas - TI e Unidades de Conservação – UC), a inexistência de licenças de operação, ao descarte inadequado de contaminantes químicos, a não execução do plano de recuperação ambiental e a criação de um mercado paralelo de comercialização ativos minerais (HASLAM; ARY TANIMOUNE, 2016; HELWEGE, 2015; URKIDI; WALTER, 2011).

Do ponto de vista espacial, a mineração não se distribui homogeneamente pelos estados brasileiros. Em 2020, três de cada quatro hectares minerados no Brasil estavam na Amazônia (SOUZA *et al.*, 2020; TASSO AZEVEDO, 2024). O bioma concentra 72,5% de toda a área minerada do país, agregando mineração industrial e garimpeira, 149.393ha. Destes, 101.100ha (67,6%) são de garimpo. A quase totalidade (93,7%) do garimpo do Brasil concentra-se na Amazônia. No caso da mineração industrial, o bioma responde por praticamente a metade (49,2%) da área ocupada por essa atividade no país (MAPBIOMAS, 2021).

A expansão garimpeira coincide com o avanço sobre territórios indígenas e unidades de conservação. Em 2022, metade da área nacional de garimpos estava em UCs (40,7%) ou em TIs (9,3%). As maiores áreas de garimpo em terras indígenas estão em território Kayapó (7602 ha) e Munduruku (1592 ha), no Pará, e Yanomami (414ha), no Amazonas e Roraima. Entre as UCs as três maiores são a APA do Tapajós (34.740ha), a Flona do Amanã, (4.150ha) e o Parna do Rio Novo (1.752ha), todas no Pará (CORTINHAS FERREIRA NETO *et al.*, 2024; TASSO AZEVEDO, 2024).

Neste contexto, a classificação digital da atividade de extração mineral por meio de imagens de satélite e técnicas de DPI é uma tarefa tradicionalmente executada, mas em escalas temporais e espaciais limitadas (DINIZ, C. G. *et al.*, 2015). Além disso, a classificação geralmente agrupa a mineração de garimpo e a industrial sob o mesmo rótulo (ASNER *et al.*, 2013; GAO, B. C., 1996; LOBO, F. DE L. *et al.*, 2018; LOBO, F. L.; COSTA; NOVO, 2015; SIMIONATO; BERTANI; OSAKO, 2021; SONTER *et al.*, 2014). Portanto, superar os desafios impostos pela quantidade, diversidade e volume de dados a serem processados para garantir o mapeamento completo da Amazônia Legal Brasileira - ALB, utilizando um intervalo de mais de três décadas de dados orbitais, é uma tarefa complexa que depende de diversas técnicas da computação e da estatística para a viabilizar a execução.

1.1 Problemática e Justificativa

A atividade de extração mineral é um importante constituinte da economia brasileira, especialmente para os estados que compõem a Amazônia Legal Brasileira (ALB) e Minas Gerais. Contudo, até o momento deste trabalho as informações sobre a atividade são apenas parcialmente reconhecidas, em geral, por projetos do governo que funcionam em outras escalas temporais e do espaço, portanto são limitadas as informações quantitativas sobre a extensão territorial destinada a essa atividade e, principalmente, sobre sua evolução ao longo do tempo, de 1985 a 2022.

A escassez de dados quantitativos sobre a extração mineral no Brasil decorre principalmente da dificuldade em separar o padrão de extração mineral expresso em uma imagem digital de alvos espectro-temporalmente similares. Algoritmos tradicionais de

aprendizado de máquina, frequentemente utilizados no sensoriamento remoto, enfrentam desafios devido à confusão espectral e temporal que caracteriza a mineração nessa região.

Avaliando o componente da resposta espectral, a mineração carrega consigo grandes similaridades com alvos que apresentem solos naturalmente expostos, como em topos de morro, cerrados e campos naturais, ou ainda com atividades humanas, que não são naturais, mas que eventualmente causem a exposição do solo, como nas atividades agrícolas ou pastoris, que podem apresentar diferentes níveis de exposição do solo dependendo de sua fase de desenvolvimento (preparo do solo, semeadura, maturação vegetal e colheita).

No que diz respeito à persistência temporal, é natural existir um conjunto diversificado de consolidação desta atividade, dado o tradicionalismo da atividade na ALB, com a existência de infraestruturas mais antigas, mais recentes e padrões de abandono da atividade. Desta maneira, o domínio temporal pode renunciar à capacidade de distinguir entre padrões de extração mineral, uma vez que estes podem apresentar uma grande diversidade de comportamentos temporais, resultado das diferentes escalas de operação, nível de consolidação da atividade e eventual abandono dela.

Sendo assim, é no âmbito do contexto/espço, seja pela proximidade ao curso de rios, quase sempre dendríticos e sinuosos, ou pela formação de cavas, em geral, agrupadas e justapostas (buracos, vales e depressões são subprodutos espaciais da mineração) que se pode inferir sobre o rótulo da classe de mineração. Nesse sentido, a aplicação de técnicas de *Deep Learning* habilita a exploração do domínio contextual do sensoriamento remoto e possibilita a construção de metodologias automáticas e que generalizam a detecção da mineração, reduzindo a necessidade de retreinamento ao mudar a região de interesse.

Uma das principais contribuições deste trabalho é a implementação de um pipeline de processamento híbrido. Esse pipeline inovador combina a escalabilidade e a flexibilidade da computação em nuvem para o processamento distribuído de grandes volumes de dados orbitais com o poder computacional de uma única GPU local para a execução eficiente de tarefas específicas de *deep learning*. Essa estratégia permite superar a dependência de infraestruturas de hardware de alto custo, como clusters computacionais e múltiplas GPUs, tradicionalmente

exigidas por classificadores baseados em arquiteturas de U¹ (BAO *et al.*, 2021; JIAO *et al.*, 2020, 2021; ZHOU *et al.*, 2018). Adicionalmente, o formato TFRecord (ABADI *et al.*, 2016; MURRAY *et al.*, 2021) é empregado para otimizar o armazenamento e a recuperação eficiente dos 37 anos de dados orbitais abrangendo todo o território nacional, explorando os recursos do Google Earth Engine (GORELICK *et al.*, 2017)

A computação aplicada é, portanto, ferramenta impreterível para a construção de um método de classificação anual, automatizado e robusto o suficiente de maneira a se adequar as necessidades de escala espaço-temporal de todo Brasil. Esta pesquisa cobre, portanto, um hiato de informação espaço-temporal na ciência brasileira, disponibilizando dados primários sobre a evolução dos padrões de extração mineral no Brasil de maneira anual e sistemática, desde 1985 até 2022.

Mais além, a disponibilização de dados primários desta natureza, fomentam frentes de pesquisa de domínios outros que tenham interesse, por exemplo, no reconhecimento do grau de ilegalidade da mineração principalmente na ALB e em territórios protegidos, na relação da atividade com o aumento da concentração de mercúrio em seus rios (BERZAS NEVADO *et al.*, 2010), na difusão da invasão de terras protegidas como TIs e UCs, no entendimento da logística de transporte do ouro, no aumento da siltação e assoreamento dos rios amazônicos (TELMER *et al.*, 2006) e outros tantos mais a se beneficiar da metodologia apresentada e dos achados desta tese.

1.2 Caracterização da Classe Alvo

O que nesta tese é chamado de classe alvo ou de mineração corresponde a toda e qualquer planta de atividade de extração mineral à céu aberto, independente de seu formato ou mineral a ser extraído.

Ao longo desta tese a mineração pode ser subdivida em duas categorias, mineração industrial e garimpo. A mineração industrial, ilustrada na Figura 1, corresponde para essa tese

¹ Arquitetura de U, como UNet e ResUnet, são caracterizadas por um formato arquitetural que se assemelha a um U, com caminhos de contração e expansão que permitem a combinação de informações contextuais e de localização.

a toda planta de mineração que tenha seu registro na Agência Nacional de Mineração - ANM como em “Autorização de Lavra” desde que na ANM não esteja caracterizado como garimpo.

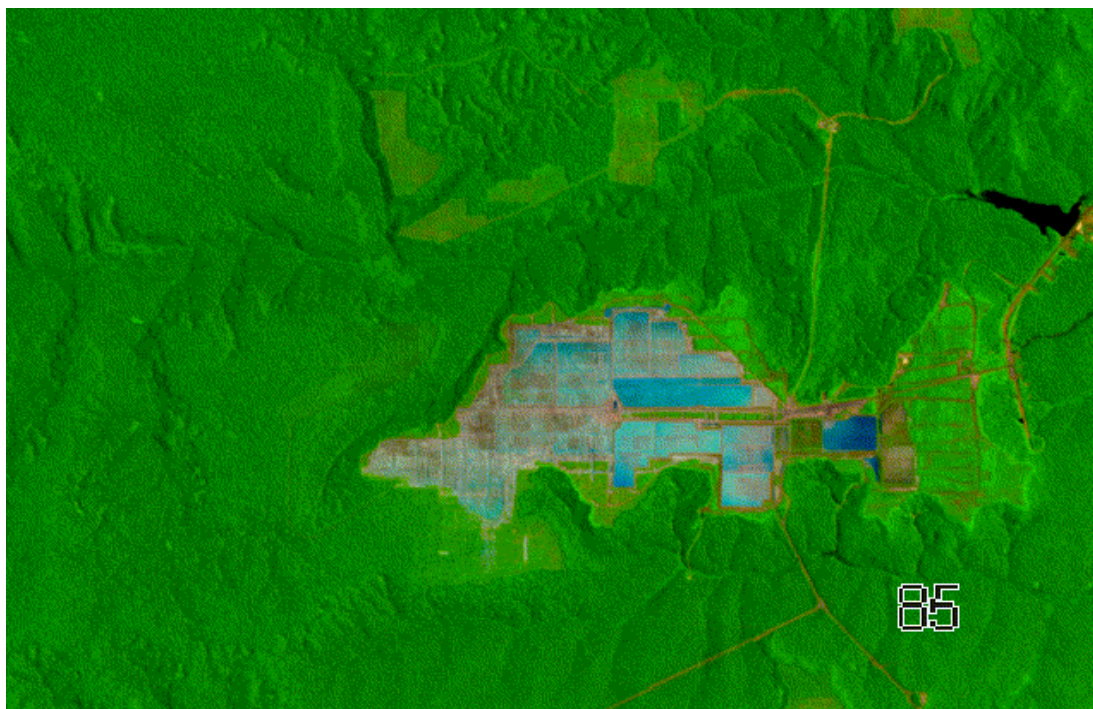


Figura 1 - Planta de Mineração Industrial - Mosaico de imagens Landsat, fonte: O autor (2025)

Os garimpos, ilustrados na Figura 2, consistem na exploração de depósitos minerais superficiais ou de subsolo de baixa profundidade, caracterizada pela utilização de técnicas e equipamentos de extração rudimentares ou de baixa complexidade tecnológica. Os garimpos têm o formato de meandros de rios, pois seguem o curso dos rios, estão fortemente ligados a extração de ouro embora metais como cobre e cobalto também sejam minoritariamente explorados (SONTER *et al.*, 2014).



Figura 2 – Mineração do tipo garimpeira - Mosaico de imagens Landsat, fonte: O autor (2025)

Garimpos e plantas de mineração industriais se diferem pela forma com que se apresentam na superfície, plantas de mineração levam em conta o rigor da engenharia e planejamento das áreas de mina, bem como as legislações vigentes para o projeto de lagos de rejeito e estruturas associadas. Os garimpos tem formato diverso, normalmente formatados são como os cursos de aluviões. Para essa classe todas as estruturas que correspondam a mina são consideradas alvo, exceto por estradas que podem por vezes estarem adjacentes.

1.3 Objetivo

O objetivo desta tese é propor uma metodologia eficaz e automática baseada em DL para detectar e mapear as mudanças de uso do solo associados a atividade de exploração mineral a céu aberto no Brasil no período de 1985 a 2022.

Destacam-se, a seguir, os objetivos específicos:

- i. Criar uma metodologia baseada em inteligência computacional, aplicada para a adequada identificação da classe de mineração.

- ii. Realizar a detecção de mudanças da mineração entre os anos especificados, analisando variações positivas e negativas de área mapeada anualmente.
- iii. Segregar a classe de mineração em: “Mineração Industrial” e “Garimpo”
- iv. Analisar espacialmente a expansão da mineração em diferentes recortes territoriais, incluindo estados, municípios, terras indígenas e unidades de conservação.

1.4 Estrutura da Tese

Esta tese, está organizada da seguinte forma:

1. Introdução: apresenta de forma breve o problema a ser contornado, e os objetivos desta tese.
2. Fundamentação Teórica: apresentando todo o arcabouço teórico necessário para o entendimento do desafio de mapeamento e importância da mineração brasileira, bem como, possui subcapítulos para explanar características da aplicação e os fundamentos das técnicas computacionais aqui utilizadas.
3. Estado da Arte de Segmentação Semântica: Posiciona este trabalho frente as novidades técnicas da segmentação semântica, comparando o modelo da U-Net original com os cinco algoritmos mais citados encontrados em revisão da literatura. O capítulo também apresenta as modificações propostas na U-Net para que ela tenha seu consumo computacional diminuído.
4. Método Proposto: aqui é explanada a forma proposta para solucionar o problema, contém informação importante sobre os desafios que se apresentaram e como foram solucionados.
5. Resultados: neste capítulo estão dispostos os resultados obtidos, e destacados, os que cumprem com os objetivos específicos, quase sempre em formato de ilustração, gráfico ou tabela para otimizar a leitura.
6. Conclusão: apresenta as conclusões obtidas pelo preenchimento da lacuna aqui discutida e a oportunidade de futuros trabalhos.

2 Fundamentação Teórica

2.1 Mineração no Brasil

A extração mineral foi uma das atividades trazidas com os Portugueses, tendo o primeiro registro da atividade mineral ainda nos primeiros anos do século 15 com a busca por metais nobres e gemas. (MACHADO; SILVA, 2001). Entretanto, a mineração se firma como atividade economicamente significativa apenas no terceiro ciclo econômico do país, também marcado pela produção de açúcar e pela expansão territorial brasileira. Trinta anos depois, é concentrada a busca por ouro de aluvião na região de Minas Gerais, e após o século 19 estende sua área de busca para região Amazônica, que nesta ocorre em simultâneo com a exploração do látex de seringais, matéria prima da borracha.

Porém, o padrão atual de exploração, se inicia em 1979. Neste ano, há um grande aumento no preço do ouro no mercado internacional, evento que efetivamente dispara a busca pelo ouro na Amazônia Brasileira, também conhecido como “corrida pelo ouro”. Com isso a mineração de ouro se torna a atividade mais economicamente representativa na região sudeste paraense. Esta época fica marcada tanto pelo crescimento do número de pequenas explorações de ouro de aluvião, quanto pelo surgimento de grandes cavas, como a de Serra Pelada (LARRETA, 2002).

Em 1983 o governo brasileiro implementa políticas para regularização da atividade mineral, e isso dá início ao declínio da atividade mineral ilegal de grande escala. Em 1990, é no governo do presidente Collor que diminui o preço da grama de ouro, tornando menos atrativa sua exploração, e é aumentado o custo dos insumos para extração do ouro, além da criação de leis que proíbem explorações minerais nas terras Indígenas, dando base a legislação mineraria vigente (MACHADO; SILVA, 2001).

Apesar do Brasil historicamente possuir um conjunto de leis que regulamentem a extração mineral, é um fato que muitas cidades ainda têm sua economia fortemente atrelada ao garimpo de ouro ilegal e de que essa atividade ainda não é plenamente fiscalizada (SONTER *et al.*, 2014).

Em 2010, o preço do ouro no mercado internacional se elevou de maneira intensa, resultando em uma alta busca pela venda de ouro e por consequência o reforço do interesse dos

garimpeiros na Amazônia Brasileira. O emergente avanço dos garimpos na ALB é um contraponto a todas as políticas e propostas de redução do desmatamento para conservação da vida selvagem e regulação climática (IUBEL, 2020; SIMIONATO; BERTANI; OSAKO, 2021).

Após estes 524 anos de convivência com a atividade mineral, hoje tentamos proteger a floresta nativa unindo a força das leis às técnicas de mapeamento e fiscalização, por isso, somos capazes de separar a mineração em duas categorias quanto ao seu tipo de implantação:

1. Legais, explorações minerais com lavra autorizada pelo órgão competente, aqui existe uma importante compensação econômica e ambiental para cada extensão minerada.
2. Ilegais, sem autorização por órgão responsável, normalmente implantada em área protegida ou de ocupação restrita, marcada pelo extrativismo, causando dano ambiental sem nenhuma regra e sem nenhum retorno econômico/social para o país (SONTER *et al.*, 2017).

Conscientemente, o Brasil criou artifícios para mitigar e fiscalizar atividades ilegais de mineração: como a criação de um grande projeto de detecção diária que é o DETER-B do instituto nacional de pesquisa espacial – INPE, que busca alertar em tempo real danos causados à floresta, dentre eles a mineração (DINIZ, C. G. *et al.*, 2015).

2.2 Sensoriamento Remoto e seus Domínios

O sensoriamento remoto obtém informações sobre a superfície da Terra a partir de três domínios fundamentais (JENSEN, J. R., 2009; PONZONI; SHIMABUKURO, 2009): espectral, espacial e temporal, conforme ilustrado na Figura 3. O domínio espectral refere-se ao comprimento de onda ou frequência e às propriedades reflexivas ou emissivas. O domínio espacial abrange a forma, tamanho, posição, localização, distribuição, textura e vizinhança do objeto. Por fim, o domínio temporal considera as mudanças do objeto ou da superfície ao longo do tempo.

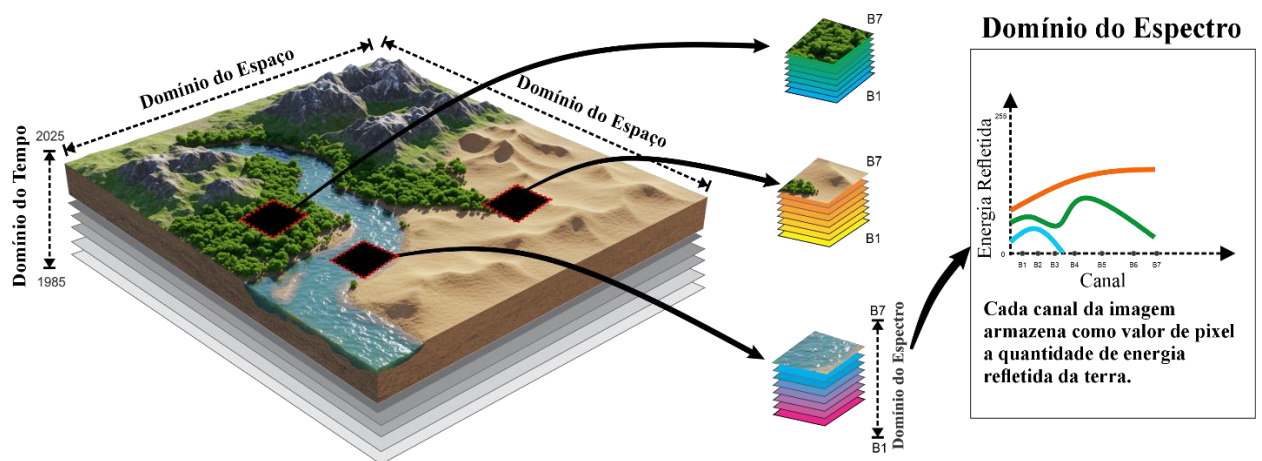


Figura 3 - Os três principais domínios de informação do sensoriamento remoto. Os domínios do espectro, do tempo e do espaço. Fonte: O autor (2025).

A maioria das pesquisas em sensoriamento remoto busca estabelecer uma relação determinística entre as propriedades químicas, biológicas e físicas dos fenômenos, superfícies ou objetos estudados e a quantidade de energia eletromagnética que eles refletem, emitem ou espalham (DE MORAES NOVO, 2010). Neste contexto, o domínio espectral fornece a informação crucial para a distinção entre alvos, como "A" e "B", na superfície terrestre. A resolução espectral, por sua vez, é definida pelo número e pela largura das faixas de comprimento de onda (bandas) às quais um sensor é sensível (JENSEN, JOHN R., 2014).

No entanto, características exclusivamente espectrais podem não ser suficientes para diferenciar dois ou mais alvos. Nestes casos de similaridade espectral, é preciso recorrer a atributos de outros domínios, como o temporal ou o espacial, para individualizar os alvos. Isso ocorre, por exemplo, na distinção entre pastagens e certos tipos de culturas agrícolas (ADAMI *et al.*, 2012; RUDORFF *et al.*, 2012).

Esses dois usos da terra podem apresentar sinais espectrais característicos de diversos momentos fenológicos². Pastagens e canaviais, por exemplo, embora sejam usos da terra distintos, passam por fases fenológicas similares e, conseqüentemente, exibem diferentes padrões espectrais ao longo do seu desenvolvimento. Contudo, ambos podem apresentar

² Fenológico - Refere-se ao estudo dos fenômenos periódicos dos seres vivos e suas relações com as condições do ambiente e a correlação com os aspectos morfológicos.

simultaneidade temporal entre uma ou mais fases de seu desenvolvimento (preparo do solo, semeadura, maturação vegetal e colheita). Assim, sem o auxílio de séries temporais, sempre haveria a probabilidade de se observar, em um momento instantâneo, a mesma fase de desenvolvimento fenológico, tanto em pastagens quanto em culturas agrícolas. Portanto, o domínio temporal é crucial para a separação entre padrões agrícolas e pastagens, como demonstrado na Figura 4.

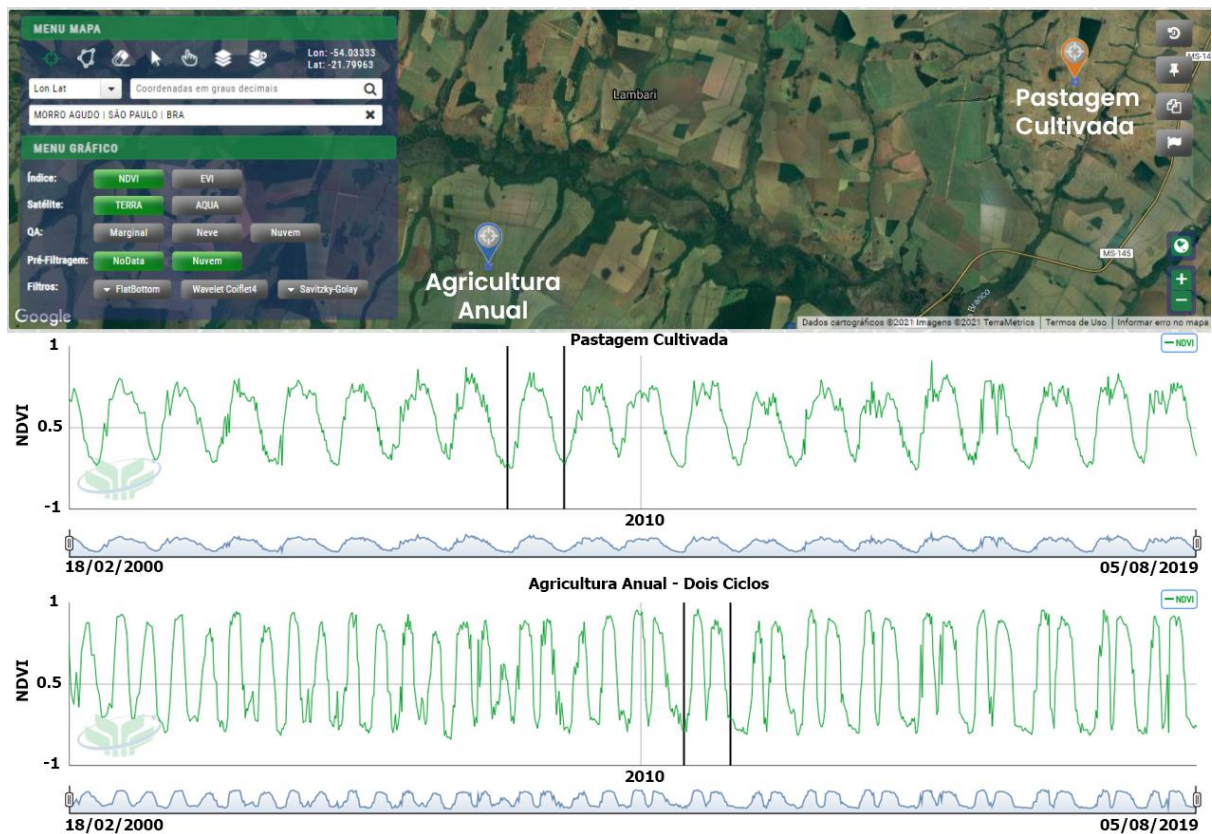


Figura 4 - Representação da Plataforma SatVeg (EMBRAPA-CNPTIA). Variação temporal de séries NDVI-MODIS sobre pastagem cultivada e agricultura anual de dois ciclos. Fonte: (SILVA *et al.*, 2022)

Em muitos casos, nem mesmo as propriedades temporais e espectrais são suficientes para discriminar alvos "super semelhantes", que se comportam de maneira similar tanto no domínio espectral quanto no temporal. Um exemplo disso é a separação entre padrões aquícolas e salineiros de outros alvos (DINIZ, C. *et al.*, 2021), que apresentam água em superfície como característica predominante ao longo do tempo, como ilustrado na Figura 3.

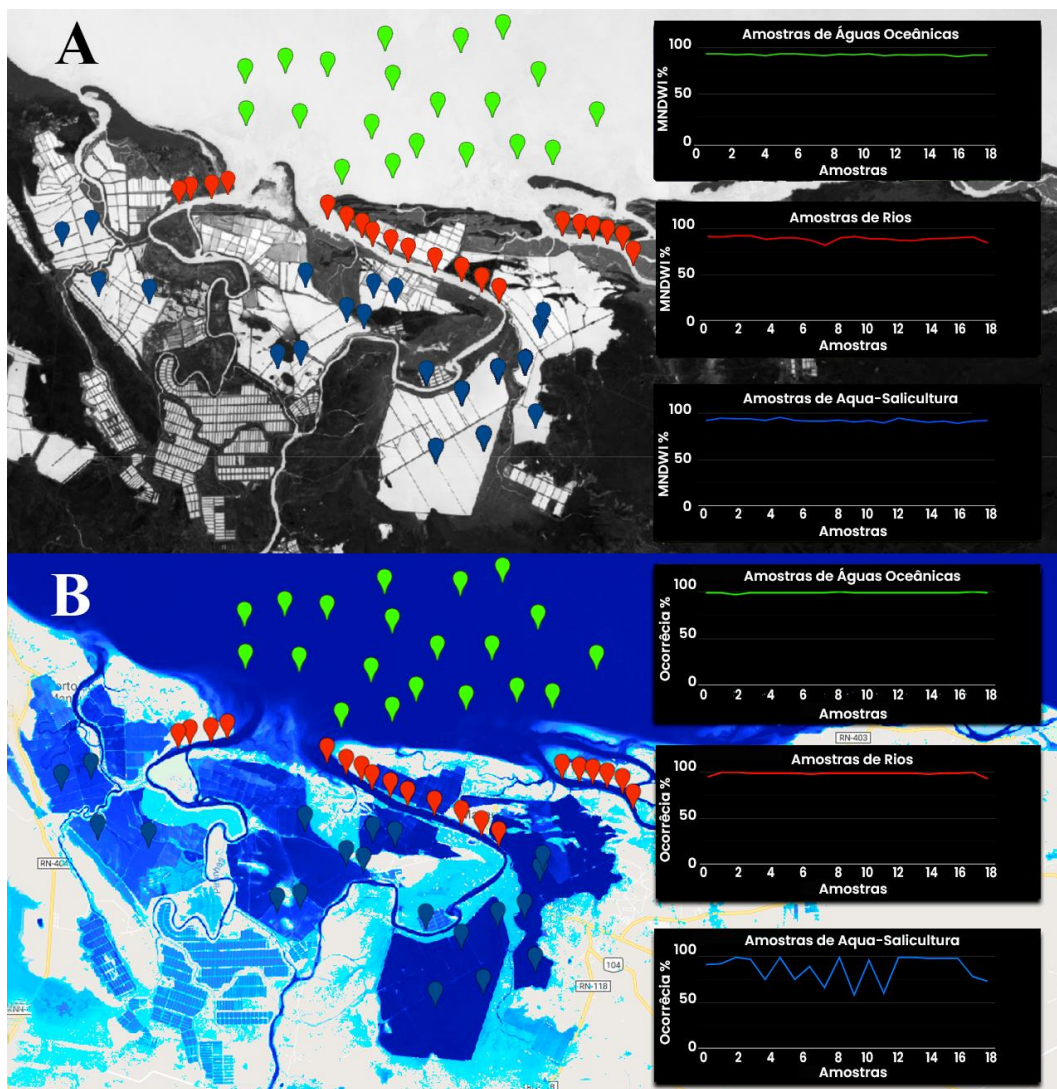


Figura 5 - Características espectrais e temporais de corpos d'água. Em A, variação do MNDWI, em três alvos distintos. Em B, percentual de ocorrência de água superficial ao longo de 34 anos. Ainda em B, a tonalidade azul escuro representa ~100% de ocorrência, azul, Fonte: (DINIZ, C. *et al.*, 2021).

Pelo disposto na Figura 3, é evidente que o domínio temporal, isoladamente, pode não ser capaz de distinguir entre padrões aquícolas/salineiros bem consolidados, canais principais de rios, lagos e lagoas perenes e águas abertas (seção B da Figura 3), uma vez que todas essas feições podem apresentar alta persistência temporal de água ao longo dos anos.

Em situações em que a cobertura natural entre os alvos é a mesma, como solo naturalmente exposto e solo exposto por cavas de mineração, e a persistência temporal também é similar, o domínio espacial/contexto, ao invés do espectral ou temporal, pode ser a chave para a distinção.

A análise da vizinhança é fundamental para acessar o componente espacial/contexto. Neste cenário, os classificadores que possuem a capacidade de reconhecer seus alvos considerando o contexto levam vantagem na previsão do rótulo de cada pixel, partindo de regiões de treinamento ao redor de cada posição da matriz. A estratégia de agrupamento da vizinhança, de diferentes maneiras, fundamenta o funcionamento dos classificadores contextuais, sejam eles orientados a objeto – OBIA (BAATZ; SCHÄPE, 2000); baseados em redes neurais convolucionais - CNN (FU, Y. *et al.*, 2019); ou no uso combinado dos dois – OCNN (ZHANG, Z.; LIU; WANG, 2018); ou, mais recentemente, no uso de *transformers* para a tarefa de segmentação de imagens de satélite (WANG, RUIKUN *et al.*, 2024). Todas essas três técnicas oferecem a capacidade de acessar o "domínio do contexto" ao invés de se ater a valores de pixels isolados.

2.3 Programa Landsat

O programa Landsat (WULDER *et al.*, 2022), iniciado no começo da década de 1970, revolucionou o sensoriamento remoto da Terra. Com sete missões bem-sucedidas ao longo de mais de 40 anos, o programa documentou e continua a documentar a superfície terrestre e suas transformações. As missões e sensores Landsat evoluíram junto com a tecnologia, transitando de um projeto inicialmente atrelado à interpretação visual analógica para uma missão totalmente digital, em escala global, com aprimoramentos contínuos em termos de desempenho espectral, espacial, radiométrico e geométrico, além de melhorias na estratégia de aquisição e disponibilidade de dados e produtos (EMERY; CAMPS, 2017).

A série de satélites Landsat constitui o mais longo registro contínuo de observações da Terra. Como tal, representa um recurso inestimável para o monitoramento de mudanças globais, sendo uma fonte primária de observações da Terra em média resolução e, consequentemente, uma ferramenta fundamental para o auxílio à tomada de decisões (CHANDER; MARKHAM; HELDER, 2009; MASEK *et al.*, 2020)

O Programa Landsat teve início em 1972 com o lançamento do primeiro satélite da série, originalmente chamado de *Earth Resources Technology Satellite* (ERTS-1) e posteriormente renomeado para Landsat 1. Com o avanço das capacidades tecnológicas, a quantidade e a qualidade dos dados capturados pelos diversos sensores a bordo dos satélites também

aumentaram. A Tabela 1 e Figura 6, apresentam informações gerais sobre cada um dos satélites Landsat.

Tabela 1 - Informações gerais sobre cada um dos satélites.

Satélite	Sensores	Lançamento	Desativação	Inclinação (graus)	Período (min)	Revisita
Landsat 1	MSS e RBV	23/07/1972	07/01/1978	99.20	103.34	18 dias
Landsat 2	MSS e RBV	22/01/1975	25/02/1982	99.20	103.34	18 dias
Landsat 3	MSS e RBV	05/03/1978	31/03/1983	99.20	103.34	18 dias
Landsat 4	MSS e TM	16/07/1982	30/06/2001	98.20	98.20	16 dias
Landsat 5	MSS e TM	01/03/1984	05/06/2013	98.20	98.20	16 dias
Landsat 6	ETM	05/10/1993	Não alcançou a órbita	---	---	---
Landsat 7	ETM+	15/04/1999	05/04/1999	98.20	98.20	16 dias
Landsat 8	OLI e TIRS	11/02/2013	---	98.20	98.20	16 dias
Landsat 9	OLI2 e TIRS2	15/02/2023	---	98.20	98.20	16 dias



Figura 6 - Data de lançamento de cada um dos satélites da família Landsat. Fonte: USGS.

Desde o início do programa, ao menos um satélite Landsat esteve em operação, assegurando mais de 40 anos de imagens contínuas da superfície terrestre, como demonstrado na Figura 6. Dos nove sistemas de satélites desenvolvidos, apenas o Landsat 6 não conseguiu

atingir sua órbita. Seu predecessor, o Landsat 5, detém o recorde de longevidade entre os satélites de imageamento terrestre, permanecendo em atividade por 29 anos.

As características espectrais dos sensores da família Landsat variaram ao longo do tempo, como ilustrado na Figura 7. O *Multispectral Scanner System* (MSS), a bordo dos Landsats 1-5, possuía quatro bandas espectrais distintas. O *Thematic Mapper* (TM), embarcado nos Landsats 4 e 5, contava com sete bandas. O *Enhanced Thematic Mapper Plus* (ETM+), do Landsat 7, possui oito bandas, enquanto os Landsats 8 e 9 possuem 11 bandas cada.

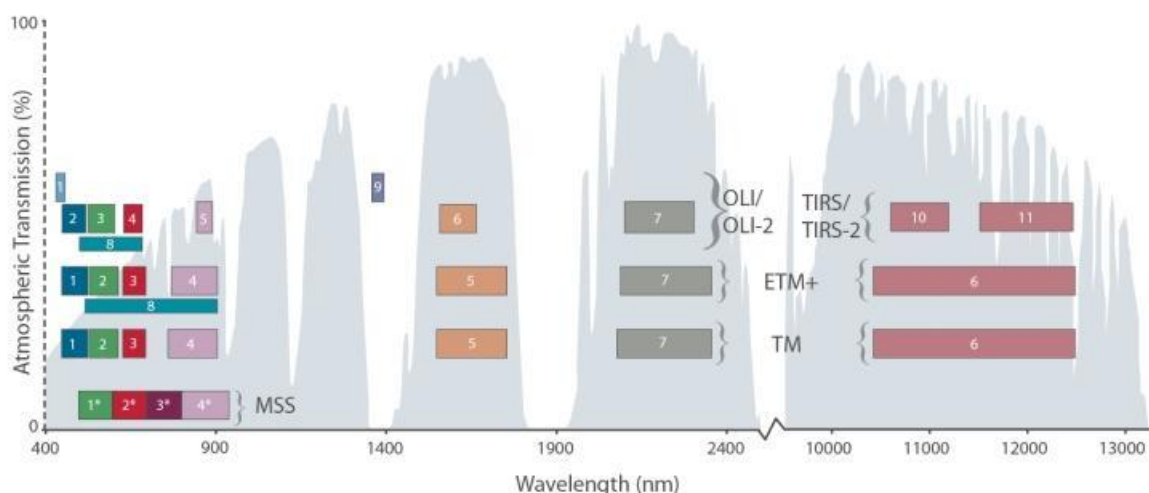


Figura 7 - Características espectrais e espaciais dos sensores da família Landsat. O Multispectral Scanner System (MSS) a bordo dos Landsats 1–5 tinha quatro bandas. O Thematic Mapper (TM) a bordo dos Landsats 4 e 5 tinha sete bandas. O Enhanced Thematic Mapper Plus (ETM+) do Landsat 7 tem 8 bandas e o Landsats 8 e 9 com 11 bandas. Fonte: USGS.

Juntas, as bandas espectrais dos diferentes satélites Landsat cobrem diversas regiões do espectro eletromagnético. A região do visível é representada pelas bandas azul (B), verde (G) e vermelho (R). A região do infravermelho é coberta pelas bandas do infravermelho próximo (NIR), do infravermelho de ondas curtas 1 e 2 (SWIR 1 e 2) e do infravermelho termal, com duas bandas nos Landsats 8 e 9 (TIRS e TIRS-2) e apenas uma nos Landsats 5 e 7.

As bandas costeira/aerossol e cirrus são inovações presentes nos Landsats 8 e 9. A banda costeira foi desenvolvida para facilitar a detecção de clorofila, um atributo importante na colorimetria de águas superficiais, especialmente em zonas costeiras (WULDER *et al.*, 2022). A banda cirrus, por sua vez, foi projetada para aprimorar a detecção deste tipo de nuvem, juntamente com suas sombras. Ela está posicionada em um intervalo do espectro eletromagnético onde as nuvens cirrus são mais visíveis.

Espacialmente, a resolução pode variar entre 15, 30, 60, 90 e 120 metros nos diferentes sensores da família Landsat, dependendo do sensor e da banda em questão. A Tabela 2 agrupa as características espaciais e espectrais dos sensores a bordo dos Landsat 5, Landsat 7 e Landsat 8.

Tabela 2 - Características espaciais e espectrais dos sensores a bordo dos Landsat 5, Landsat 7. Fonte: USGS.

Bandas Landsat-5 TM (μm)			Bandas Landsat-7 ETM+ (μm)		Bandas Landsat-8 OLI e TIRS (μm)			
---					30 m Coastal/Aerosol		0.43 - 0.45	Band 1
Band 1	30 m Blue	0.45 - 0.52	Band 1	30 m Blue	0.44 - 0.51	30 m Blue	0.45 - 0.51	Band 2
Band 2	30 m Green	0.52 - 0.60	Band 2	30 m Green	0.52 - 0.60	30 m Green	0.53 - 0.59	Band 3
Band 3	30 m Red	0.63 - 0.69	Band 3	30 m Red	0.63 - 0.69	30 m Red	0.63 - 0.67	Band 4
Band 4	30 m NIR	0.76 - 0.90	Band 4	30 m NIR	0.77 - 0.90	30 m NIR	0.85 - 0.87	Band 5
Band 5	30 m SWIR-I	1.55 - 1.75	Band 5	30 m SWIR-I	1.54 - 1.75	30 m SWIR-I	1.56 - 1.65	Band 6
Band 6	120 m TIR	10.4 - 12.5	Band 6	60 m TIR	10.3 - 12.3	100 m TIR-I	10.60 - 11.19	Band 10
						100 m TIR-2	11.50 - 12.51	Band 11
Band 7	30 m SWIR-2	2.08 - 2.35	Band 7	30 m SWIR-2	2.06 - 2.34	30 m SWIR-2	2.10 - 2.29	Band 7
---			Band 8	15 m Pan	0.51 - 0.89	15 m Pan	0.50 - 0.67	Band 8
---					30 m Cirrus		1.36 - 1.38	Band 9

2.4 Aprendizado de Máquina

Aprendizado de Máquina - AM refere-se à utilização de algoritmos para "aprender" com dados e realizar determinações sobre algo (LARY *et al.*, 2016). O AM pode ser classificado de acordo com a estratégia de extração de atributos empregada. Quando um classificador utiliza

instâncias preexistentes (amostras) para estimar, prever ou classificar padrões gerais, com o objetivo de prever ou rotular o resultado de instâncias futuras, trata-se de aprendizado supervisionado. Por outro lado, quando não há um resultado comparativo ou rótulo pré-existente antes da extração de atributos, o aprendizado é considerado não supervisionado.

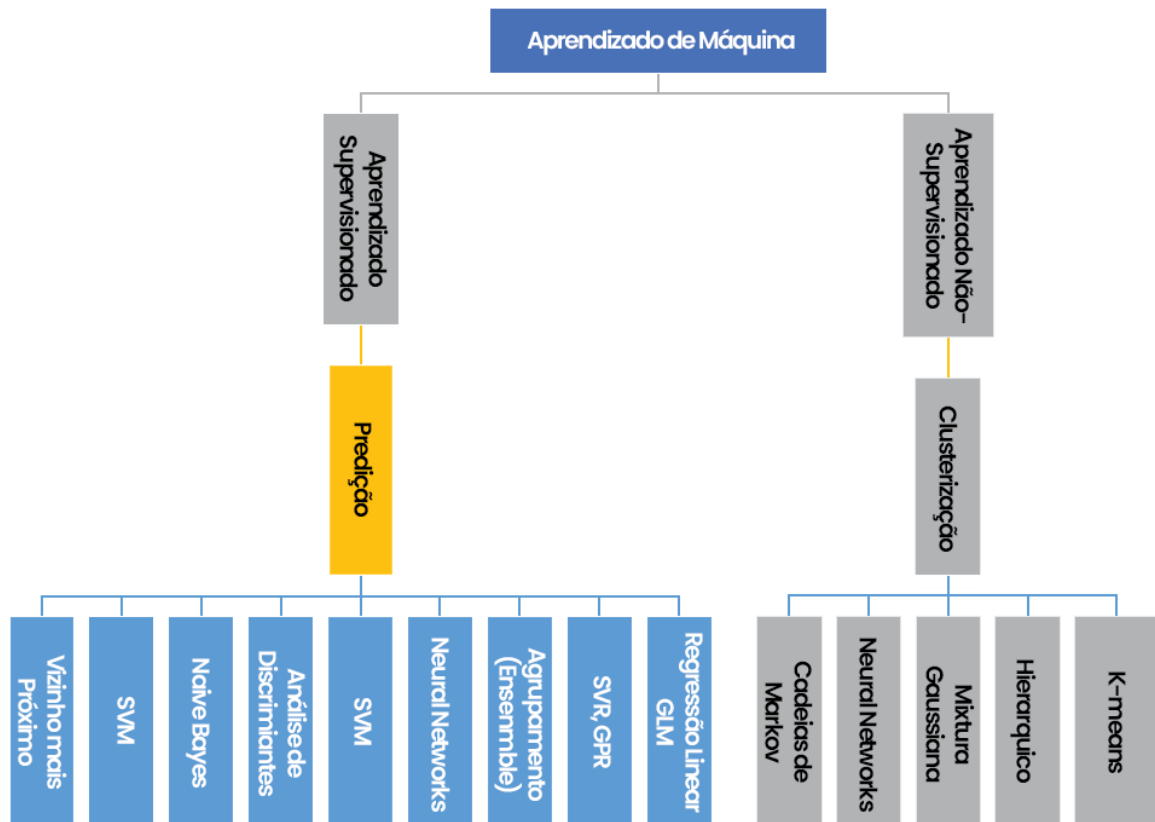


Figura 8 - O AM pode ser categorizado em supervisionado e não-supervisionado. Quando o classificador estima padrões gerais usando de instâncias preexistentes (amostras), temos o aprendizado supervisionado. Quando não há um resultado comparativo ou rótulo pré-existente, anterior a extração de atributos, considera-se o aprendizado como não-supervisionado. Fonte: O autor (2025).

Na família do aprendizado supervisionado, destaca-se o subgrupo dos algoritmos de predição, que engloba tanto os algoritmos de classificação quanto os de regressão. No ramo do aprendizado não supervisionado, os clusterizadores são os representantes mais populares Figura 8. Também Cadeias de Markov (DERRODE; PIECZYNSKI, 2016; “Markov Chains and Queuing Theory”, 2005) e redes neurais convolucionais (HINTON, G. E.; OSINDERO; TEH, 2006; YANN LECUN, YOSHUA BENGIO, 2015) têm seu uso reportado na literatura. Vale ressaltar que alguns algoritmos podem apresentar um comportamento dual, podendo ser utilizados como classificadores ou regressores, dependendo de sua parametrização. É o caso, por exemplo, do Random Forest (BREIMAN, 1999), um algoritmo de AM baseado em métodos

de agrupamento (*Ensemble Methods*) que pode funcionar tanto como regressor quanto como classificador.

Algoritmos supervisionados e não supervisionados podem ser combinados por diversos motivos e de várias formas. Neste trabalho, utilizaremos uma abordagem que une o K-means (BOCK, 2007; FORGY, 1965) um clusterizador não supervisionado, e a U-Net (RONNEBERGER; FISCHER; BROX, 2015), um classificador supervisionado. Essa combinação sequencial tem como objetivos: a) reduzir a dimensionalidade do conjunto de dados original e b) explorar o domínio espacial/contexto como forma de distinguir alvos que são extremamente semelhantes nos domínios espectral e temporal.

2.5 Clusterização com K-means

O K-means é um algoritmo de aprendizado de máquina não supervisionado, amplamente utilizado para agrupamento (*clustering*) de dados. Seu objetivo é particionar um conjunto de dados em k grupos (*clusters*) distintos, onde k é um número predefinido pelo usuário. Cada grupo é representado por seu centroide, que é a média aritmética dos pontos pertencentes ao cluster. O K-means opera de forma iterativa, buscando minimizar a variância intracluster, ou seja, a soma das distâncias quadráticas entre os pontos de dados e o centroide do seu respectivo cluster. Em termos simples, ele procura agrupar pontos de dados que são semelhantes entre si, com base em sua proximidade em um espaço multidimensional.

O funcionamento do K-means (FORGY, 1965; HARTIGAN; WONG, 1979; MACQUEEN, 1967), apresentado graficamente na Figura 9, pode ser resumido em quatro etapas principais:

- I. **Inicialização:** O algoritmo inicia com a seleção aleatória de k pontos do conjunto de dados, que servirão como os centroides iniciais dos k clusters. Este método de inicialização aleatória, embora simples, é sensível à escolha inicial e pode levar a resultados subótimos (FORGY, 1965). Métodos alternativos de inicialização, como o K-means++ (ARTHUR; VASSILVITSKII, 2007), buscam uma melhor distribuição inicial dos centróides, potencialmente levando a uma convergência mais rápida e a melhores resultados.

- II. **Atribuição:** Nesta etapa, cada ponto do conjunto de dados é atribuído ao cluster cujo centroide está mais próximo. A proximidade é tipicamente calculada usando a distância Euclidiana, que mede a distância em linha reta entre dois pontos no espaço n-dimensional. No entanto, outras métricas de distância, como a distância de Manhattan (SUWANDA; SYAHPUTRA; ZAMZAMI, 2020) ou a distância de Minkowski (SINGH; YADAV; RANA, 2013), também podem ser empregadas dependendo da natureza dos dados. A escolha da métrica de distância pode influenciar a forma dos clusters encontrados.
- III. **Atualização:** Após a atribuição de todos os pontos aos seus respectivos clusters, os centroides são recalculados. Cada novo centroide é definido como a média aritmética das coordenadas de todos os pontos pertencentes ao seu cluster.
- IV. **Iteração:** As etapas de atribuição (II) e atualização (III) são repetidas iterativamente. A cada iteração, os pontos podem ser realocados para diferentes clusters, e os centroides são ajustados de acordo com as novas atribuições. O processo iterativo continua até que um critério de convergência seja satisfeito. Os critérios de convergência mais comuns incluem: (a) a estabilização dos centroides, ou seja, quando os centroides não mudam significativamente entre iterações sucessivas, ou (b) a obtenção de um número máximo predefinido de iterações. O algoritmo também pode convergir quando a atribuição de pontos aos clusters não muda mais (JAIN, 2010).

Funcionamento do K-Means

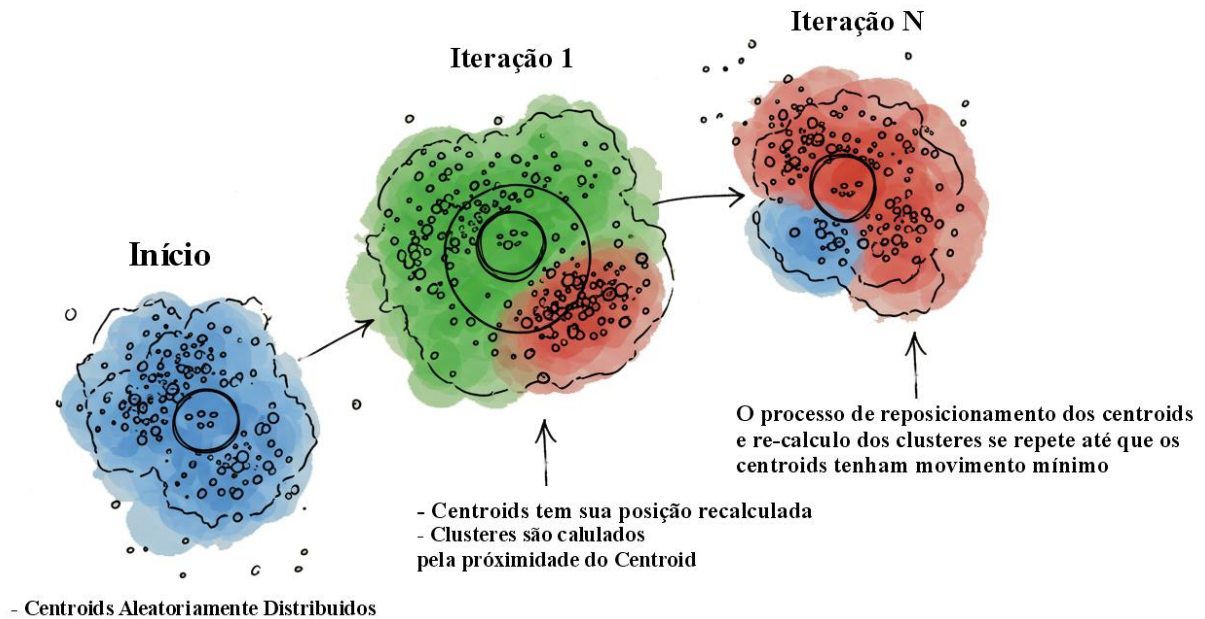


Figura 9 - O algoritmo K-means separa n pontos de dados em k clusters. De tal forma que, cada elemento pixel do dado de entrada seja classificado em um cluster que tenha a menor diferença entre o valor médio (centróide) do cluster e o valor do pixel em questão. Fonte: O autor (2025).

Diferentes inicializações podem levar a diferentes resultados de agrupamento. Para mitigar esse problema, é comum executar o algoritmo várias vezes com diferentes inicializações aleatórias e escolher a solução com a menor variância intracluster. Além disso, a escolha do valor de k é crucial e pode impactar significativamente a qualidade dos clusters. Existem métodos para auxiliar na escolha de um valor adequado para k , como o método do cotovelo (KRZANOWSKI; LAI, 1988) e a análise de silhueta (ROUSSEEUW, 1987).

O K-means é um algoritmo de agrupamento iterativo, simples e eficiente, que busca particionar um conjunto de dados em k clusters, minimizando a distância entre os pontos de dados e os centroides dos seus respectivos clusters. Apesar de sua simplicidade, o K-means é amplamente utilizado em diversas áreas (LV *et al.*, 2010), como segmentação de mercado (LI, Y. *et al.*, 2021), reconhecimento de padrões (PENG *et al.*, 2013), compressão de imagens (PAUL *et al.*, 2024) e análise de dados em geral (IKOTUN *et al.*, 2023). Sua popularidade se deve à sua facilidade de implementação e interpretação, além de sua capacidade de lidar com conjuntos de dados relativamente grandes.

2.6 Deep Learning

Deep learning (DL) é uma subárea do aprendizado de máquina (AM) que engloba algoritmos baseados em redes neurais com múltiplas camadas. Devido à sua arquitetura profunda, esses algoritmos são especialmente adequados para tarefas de classificação nas quais os alvos dependem fortemente do contexto em que estão inseridos. Alguns exemplos notáveis de aplicação de *Deep Learning* incluem: reconhecimento de objetos em imagens (KRIZHEVSKY; SUTSKEVER; GEOFFREY E., 2012), detecção de alvos por meio do comportamento de série temporal (ISMAIL FAWAZ *et al.*, 2019) e delineamento de alvos (segmentação semântica) (HAO; ZHOU; GUO, 2020). Estes são apenas alguns exemplos do uso de DL.

Segundo SCHMIDHUBER,(2015) o DL tem seu nascimento conceitual baseado em redes neurais, contemplando toda arquitetura de neurônios interligados e suas diversas possibilidades de funções de ativação. Em concordância, para LITJENS *et al.*,(2017), o DL é um aprimoramento conceitual das redes neurais clássicas, como as *Multi-Layer Perceptrons* (MLP) (ROSENBLATT, 1958; RUMELHART; HINTON; WILLIAMS, 1986), pois agrega a ideia de múltiplas camadas e ferramentas matemáticas como convoluções e interpolações. De fato, diversas referências ressaltam que as Convolutional Neural Networks (CNNs) (LECUN *et al.*, 1989, 1998) foram importantes para que a comunidade científica investisse nas Deep Neural Networks (DNNs) (GERS; SCHMIDHUBER; CUMMINS, 1999; KRIZHEVSKY; SUTSKEVER; GEOFFREY E., 2012; NAIR; HINTON, 2008) para a detecção de objetos em imagens, como a AlexNet (KRIZHEVSKY; SUTSKEVER; GEOFFREY E., 2012).

Computacionalmente, o DL apresenta um desafio: o recurso computacional necessário para o processo de treinamento de suas múltiplas camadas nas mais diversas arquiteturas. O uso de recursos é maior quando comparado com algoritmos de machine learning supervisionados, como SVM (*Support Vector Machines*) (HUVAL *et al.*, 2015) e *Random Forest* (BOSTON *et al.*, 2022). Geralmente, pelo consumo computacional, algoritmos de DL tendem a ter menor prioridade, uma vez que o recurso computacional é proporcional ao investimento financeiro. Devido à necessidade de hardware específico, frameworks e fabricantes de hardware promovem o uso de arquiteturas projetadas para o processamento paralelo em alta velocidade,

como as GPUs, para acelerar processos de treinamento, teste e predição das técnicas de DL. Exemplos incluem: NVIDIA CUDA, Intel Xen Phi e soluções em FPGA (WANG, C. *et al.*, 2017).

2.7 *Deep Learning* aplicado ao Sensoriamento Remoto

Sensoriamento Remoto é definido pelo uso de técnicas sem contato físico para medir, mensurar ou reconhecer determinado alvo (DE MORAES NOVO, 2010). Satélites e drones, ao capturarem imagens da superfície terrestre, utilizam-se do sensoriamento remoto. Nesse contexto, técnicas de *Deep Learning*, que se destacam pelo reconhecimento e/ou classificação de objetos em imagens, têm se tornado uma ferramenta para o sensoriamento remoto (MA, L. *et al.*, 2019).

As Redes Neurais Convolucionais são, de longe, a técnica de DL mais aplicada ao sensoriamento remoto (ZHU *et al.*, 2017). Isso se deve à sua adaptabilidade ao conceito de multibandas, presente em imagens de drones e satélites. De estrutura simplificada, as CNNs são compostas basicamente por: I - Camadas de convolução, responsáveis por criar a generalização do conhecimento por meio de Feature Maps; II - Camadas de pooling, que correlacionam os pixels através da re-amostragem do tamanho do pixel; e III - Camadas totalmente interconectadas, que armazenam o conhecimento. Esses três tipos de camadas combinam simplicidade estatística e baixa recursividade, construindo uma maneira eficiente de detectar alvos em imagens. Entretanto, as CNNs são classificadores baseados em pixel, que interagem com a vizinhança. Sozinha, sua estrutura não consegue agregar a informação temporal presente em uma série temporal de imagens de satélite.

As *Recurrent Neural Networks* (RNNs) (BENGIO; SIMARD; FRASCONI, 1994; ELMAN, 1990) representam um subconjunto de DL responsável por agregar o comportamento temporal do alvo em séries temporais de imagens. Nas RNNs, para dados espaciais, os inputs são vetores ou listas de imagens ordenadas pela data de captura. Nessas redes, um conjunto especial de neurônios, chamado *Long Short-Term Memory Cell* (LSTM) (VAN HOUDT; MOSQUERA; NÁPOLES, 2020), é utilizado para armazenar a generalização da percepção temporal entre as imagens de entrada.

Deep Autoencoders, ou autoencoders profundos, têm sido utilizados para extração de características em imagens hiperespectrais (ZABALZA *et al.*, 2016). Essas imagens, capturadas por sensores hiperespectrais, possuem centenas de bandas espectrais de largura estreita. Comparando um satélite multiespectral, como o Landsat-8, com um satélite hiperespectral, como o Hyperion, nota-se um aumento de mais de 200 bandas. Esse salto na complexidade dos dados geraria custos computacionais elevados se fossem utilizadas apenas as redes neurais convolucionais densas. O desafio de trabalhar com entradas de alta dimensionalidade faz com que os Autoencoders se encaixem perfeitamente nesse cenário. Essas redes foram concebidas essencialmente para compactar e simplificar os dados de entrada, traduzindo suas principais características para a saída. Assim, os Autoencoders atuam como redutores de dimensionalidade, simplificando a entrada para que, em seguida, seja processada por outras redes, como as CNNs.

Generative Adversarial Networks (GANs) (LI, C. *et al.*, 2017) são arquiteturas de redes não supervisionadas, compostas por duas redes internas que se autoavaliam: 1 - Geradora, que recebe o input de dados e o modifica para "enganar" o teste de reconhecimento da rede 2; 2 - Discriminadora, que tem como objetivo julgar se o input fornecido pela Geradora é original do domínio ou uma tentativa de falsificação. No contexto do sensoriamento remoto, as GANs têm sido utilizadas para: prever séries temporais, obter super-resolução a partir de imagens de resolução mediana (LEI; SHI; ZOU, 2020) ou até mesmo para remoção de nuvens finas (ZHAO *et al.*, 2021).

2.8 Segmentação Semântica

A segmentação semântica, um processo fundamental no campo da visão computacional, tem como objetivo classificar ou rotular cada pixel de uma imagem, atribuindo a cada um deles uma categoria semântica específica (HAO; ZHOU; GUO, 2020; MO *et al.*, 2022). Em essência, essa técnica busca compreender a imagem em um nível granular, identificando não apenas a presença de objetos, mas também suas formas, contornos e relações espaciais. Diferentemente da classificação de imagens, que atribui um único rótulo a toda a imagem, ou da detecção de objetos, que localiza objetos por meio de caixas delimitadoras, a segmentação semântica fornece uma descrição detalhada do conteúdo da imagem, pixel a pixel. Essa capacidade de mapear cada pixel para uma classe específica a torna uma ferramenta em diversas aplicações.

A capacidade da segmentação semântica de fornecer informações detalhadas ao nível de pixel abre portas para uma ampla gama de aplicações no mundo real. Na área médica, por exemplo, essa técnica auxilia na detecção de doenças em imagens médicas, como tomografias e ressonâncias magnéticas, permitindo a identificação precisa de tumores, lesões e outras anomalias (IQBAL *et al.*, 2023). No contexto de veículos autônomos, a segmentação semântica é crucial para a percepção do ambiente, permitindo que o veículo identifique e compreenda elementos como faixas de pedestres, semáforos, placas de trânsito e outros veículos (GAUTAM; MATHURIA; MEENA, 2022). Além disso, a segmentação semântica tem se mostrado valiosa em aplicações de sensoriamento remoto, como a detecção de deslizamentos de terra (LU *et al.*, 2020) e o monitoramento de vazamentos de óleo em plataformas petrolíferas (CANTORNA *et al.*, 2019), contribuindo para a gestão de desastres naturais e a proteção ambiental.

A importância da segmentação semântica reside na sua capacidade de extrair informações precisas e detalhadas sobre o posicionamento e a relação entre os diferentes elementos presentes em uma imagem. Ao classificar cada pixel e considerar seu contexto espacial, a segmentação semântica permite uma compreensão mais profunda da cena, indo além da simples detecção ou classificação de objetos. Essa compreensão aprimorada é crucial para tarefas que exigem uma análise minuciosa da imagem, como a navegação autônoma (HUANG; CHEN, 2020), o diagnóstico médico assistido por computador (LITJENS *et al.*, 2017; ZHOU, T.; RUAN; CANU, 2019; ZHOU, Z. *et al.*, 2018) e a análise de imagens de satélite para monitoramento ambiental (CORTINHAS FERREIRA NETO *et al.*, 2024; DINIZ, C. *et al.*, 2021; SOUZA *et al.*, 2020). Em suma, a segmentação semântica desempenha um papel fundamental na construção de sistemas de visão computacional mais inteligentes e capazes de interagir com o mundo de forma a delinear melhor os alvos.

3 Arquiteturas de Deep Learning para Segmentação Semântica: Estado da Arte e Critérios de Escolha

A segmentação semântica, uma tarefa primordial no campo da visão computacional, dedica-se a atribuir um rótulo de classe a cada pixel em uma imagem, possibilitando a classificação e o delineamento preciso de objetos e regiões (BRAR *et al.*, 2025). Em contraste com a classificação de imagens, que categoriza a cena como um todo, ou a detecção de objetos, que delimita instâncias com caixas envolventes, a segmentação semântica oferece uma compreensão granular do conteúdo visual, fornecendo uma máscara detalhada para cada categoria de interesse. Essa capacidade de analisar e rotular em nível de pixel é fundamental para diversas aplicações que demandam a identificação e o mapeamento de elementos com precisão espacial, incluindo análises detalhadas em imagens de sensoriamento remoto (LI, J. *et al.*, 2024), tópico central desta tese.

Ampliando a discussão sobre a relevância da segmentação semântica para a análise detalhada de imagens, especialmente no contexto do sensoriamento remoto abordado por esta tese, as próximas seções aprofundam-se nos aspectos cruciais para a execução da tarefa. Inicialmente, será apresentado o estado da arte dos modelos de segmentação semântica, categorizando as principais abordagens. Em seguida, exploraremos as aplicações específicas da segmentação semântica no sensoriamento remoto, destacando seus desafios e contribuições. Finalmente, será detalhado o processo de seleção dos modelos que foram escolhidos para comparação experimental neste estudo, com base nos critérios preestabelecidos.

3.1 Estado da Arte de Modelos de Segmentação Semântica

Conforme apresentado na Seção 2.8, a segmentação semântica é crucial em diversas áreas, e a crescente demanda por soluções impulsiona o desenvolvimento de novos modelos de DL. Para compreender a variedade desses modelos, autores como (BRAR *et al.*, 2025; YU *et al.*, 2023) revisaram a literatura, identificando mais de 50 modelos distintos.

Os trabalhos de Brar *et al.*, (2025; Emek Soylu *et al.*, 2023; Wang *et al.*, 2024) fazem análise sistemática da literatura de segmentação semântica e EMEK SOYLU *et al.*, (2023)

sugere ainda uma organização hierárquica desses modelos, com base na supervisão e no método de classificação.

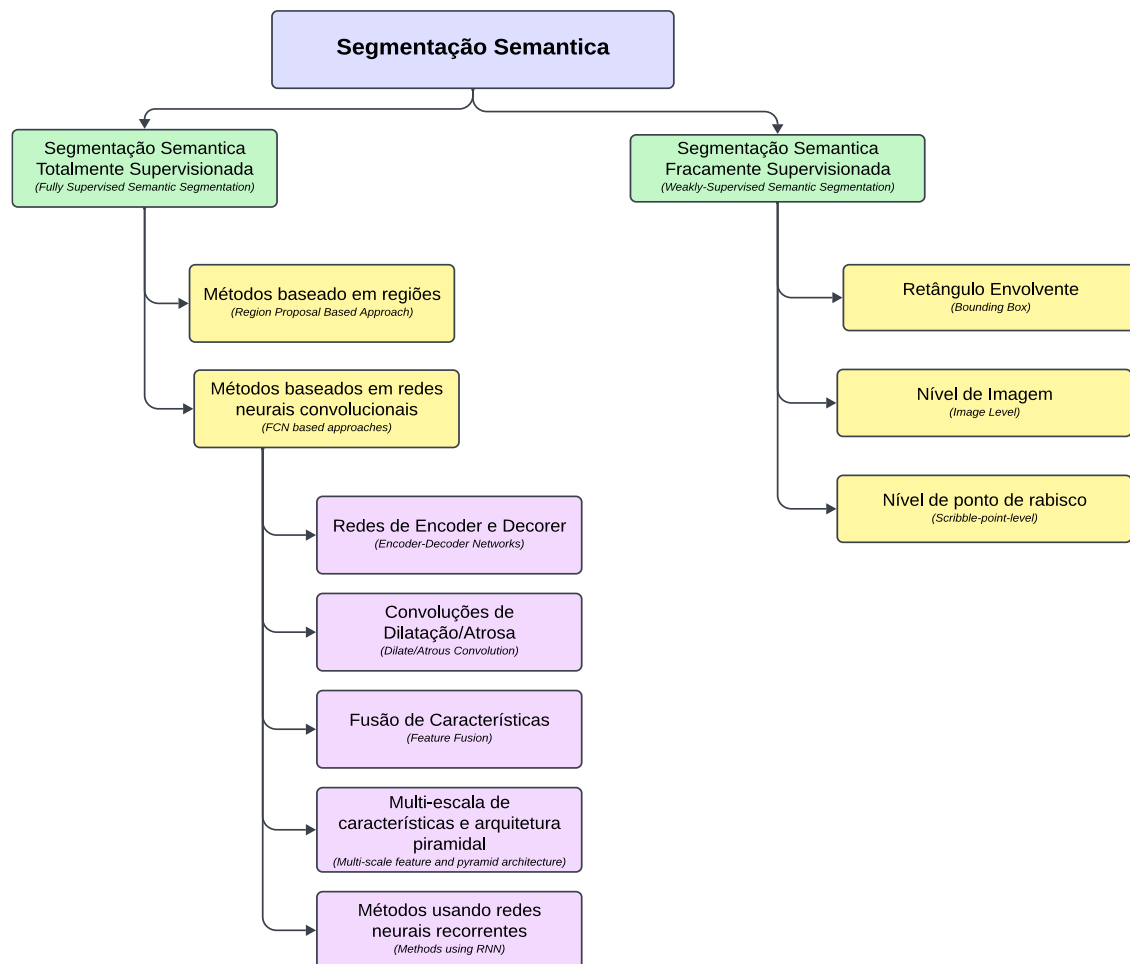


Figura 10 - Subdivisão hierárquica dos algoritmos baseados em redes convolucionais. Fonte: O autor (2025), traduzido e ajustado de (EMEK SOYLU *et al.*, 2023).

No nível mais alto, em verde, a Figura 10 apresenta dois grandes grupos de algoritmos de segmentação semântica, classificados pela forma de aprendizado:

- I. *Fully-Supervised Semantic Segmentation* (FSSS): Algoritmos que dependem de intensa supervisão humana para criar rótulos precisos em nível de pixel. Isso significa que cada pixel da imagem precisa ser rotulado para treinar o modelo (EBRAHIMI; KUMAR, 2025; WANG, RAY; CHOWDHURY; ORTIZ, 2025).
- II. *Weakly-Supervised Semantic Segmentation* (WSSS): Em contraste com a categoria anterior, esta abrange algoritmos que aprendem com menor grau de supervisão, utilizando

informações de nível superior, como rótulos de objetos ou classificações de imagens (CAI *et al.*, 2024; MEIRA, 2024; YANG, J. *et al.*, 2025). Em vez de rotular cada pixel individualmente, pode-se fornecer ao modelo informações mais gerais, como a presença de um gato na imagem, ou empregar um detector de objetos pré-treinado. Essa estratégia assemelha-se ao processo de ensino de uma criança, onde a simples exposição a um gato, sem a necessidade de descrever minuciosamente suas características, pode ser suficiente para o aprendizado.

O foco desta tese reside nos algoritmos do grupo FSSS, visto que o objetivo principal é gerar um mapa de segmentação no qual cada pixel seja classificado.

Dentro do grupo de FSSS, estão os subgrupos: *Region-Proposal-Based Approaches* (WANG, Y. *et al.*, 2025; YE; ZHU; CAO, 2023), abordagens baseadas em propostas de região, e “*FCNN Based Approaches*” em tradução abordagens baseadas em redes neurais convolucionais completamente conectadas.

A ideia principal do *Region-Proposal-Based Approaches* é gerar uma classificação em duas etapas, primeiro fazendo *Object Detection* e selecionando os *bounds* para depois delinear o alvo.

A ideia principal dos *FCNN Based Approaches* é de aplicar métodos de segmentação semântica que utilizam redes neurais convolucionais adaptadas para processar imagens, gerando mapas de segmentação densos, ou seja, com a mesma resolução da imagem de entrada, este subgrupo possui as características necessárias para o processamento de imagens de satélite é possui mais cinco divisões que aqui são brevemente caracterizadas:

- I. As redes na categoria “*Encoder-Decoder Networks*” (BAĞCI; ACAR; TÜRK, 2023; TOLAN *et al.*, 2024), são um tipo de arquitetura de rede neural, onde a rede pode ser subdividida em duas partes o *encoder* e o *decoder*, o caminho/parte de *encoder* também é chamado de “caminho de compressão” que permite a redução de dimensionalidade, extração de características, e a representação semântica. E o *decoder* ou “caminho de expansão” é o ramo final da rede que é responsável transformar a informação abstrata e compactada do *encoder* em uma saída visualmente coerente e espacialmente estruturada.
- II. As Convoluções Dilatadas (DIXIT; CHAURASIA; KUMAR MISHRA, 2021; LI, Q. *et al.*, 2024), também conhecidas como “*Dilated Atrous Convolutions*”, são um tipo de CNN que

se destaca pela capacidade de modificar os parâmetros da convolução. Isso é feito através do aumento do tamanho do kernel, da alteração do passo da convolução ou da introdução de espaçamentos (ou "buracos") entre os elementos do kernel. Essas modificações podem ser realizadas de forma dinâmica (GAO, R., 2023). A principal vantagem desse grupo de redes é a possibilidade de ajustar o campo receptivo, permitindo que a rede capture informações em diferentes escalas, sem a necessidade de alterar o número de parâmetros treináveis. Isso possibilita um controle mais granular sobre o nível de detalhamento extraído do dado de entrada.

- III. As redes com a característica de *Feature Fusion* (NYLÉN; CALLE; GONZALES-INCA, 2025), são um tipo de arquitetura de rede neural que combina informações de diferentes fontes e níveis de representação para melhorar o desempenho em tarefas de aprendizado de máquina. A ideia principal é que a combinação de diferentes tipos de características pode levar a uma representação mais rica e informativa dos dados, permitindo que a rede aprenda padrões mais complexos e faça previsões mais precisas. Existem vários mecanismos que podem suportar as redes de Feature Fusion, são eles: concatenação, soma, multiplicação e de atenção (ZHANG, Y. *et al.*, 2021). Os mecanismos podem ter variações na posição de uso na arquitetura da rede, são elas: *early fusion* (BESERRA; GOULARTE, 2023), *late fusion* (FERRARI, F. *et al.*, 2023) e *hybrid fusion* (BALAGURUNATHAN *et al.*, 2024).
- IV. As redes neurais que fazem uso da arquitetura *Multi Scale feature and pyramid* baseiam-se na premissa de que as imagens de input podem aparecer em diferentes resoluções e por isso o processo de feature extraction precisa ocorrer em diferentes tamanhos. Essas arquiteturas também fazem uso de técnicas como: *dilated convolution*, *pooling* e *upsampling*.
- V. *Methods Using RNN*, são redes que se apoiam na recursividade para a promoção de conhecimentos e entendimento das features, essas redes são vantajosas quando comparadas às CNNs por não possuírem filtros de tamanho fixo (ZHANG, F. *et al.*, 2024) o que limitaria o aprendizado a considerar aspectos temporais entre os inputs.

Além da categorização e ponderações feita por (BRAR *et al.*, 2025; EMEK SOYLU *et al.*, 2023), alguns avanços em pesquisas e modelos que se apoiam em mecanismo de atenção, dentre eles os *transformers* visuais como o ViT (KARIM *et al.*, 2025; TRAN *et al.*, 2025), demonstram potencial para classificação ao nível de pixel quando são encadeados a outros algoritmos do tipo segmentadores semânticos, a exemplo de: UNetFormer (HATAMIZADEH

et al., 2022), MaxViT-Unet (LI, X. *et al.*, 2023), Segformer (XIE *et al.*, 2021) e Segmenter (STRUDEL *et al.*, 2021).

Os modelos de estado da arte em segmentação semântica representam avanços importantes nas aplicações de DL no sensoriamento remoto, por exemplo habilitando o reconhecimento de novos alvos (TRAN *et al.*, 2025; WANG, RUIKUN *et al.*, 2024), contudo os modelos recentes baseados em *transformers* como os descritos anteriormente possuem maior custo computacional para treinamento e inferência em comparação com técnicas de segmentação semântica que não utilizem de *transformers*. Por isso, nas próximas sessões são estabelecidos os critérios de seleção do modelo a ser utilizado nesta tese.

3.2 Segmentação Semântica aplicada ao Sensoriamento Remoto

A segmentação semântica tem se tornado uma ferramenta na análise de imagens de sensoriamento remoto, permitindo extrair informações detalhadas e precisas sobre a superfície terrestre. Esta seção explora a aplicação da segmentação semântica, nesse contexto, abordando seus desafios, benefícios e as principais técnicas utilizadas.

O sensoriamento remoto fornece imagens da Terra capturadas por sensores a bordo de satélites, aviões e drones. Essas imagens contêm uma riqueza de informações sobre diferentes aspectos da superfície terrestre, como cobertura vegetal, uso da terra, corpos d'água e infraestrutura urbana. A segmentação semântica, ao classificar cada pixel da imagem em uma categoria específica, permite interpretar e analisar essas informações de forma automatizada e eficiente.

Na literatura, de acordo com recentes revisões (Adegun, Viriri e Tapamo, 2023; Li *et al.*, 2024; Yuan, Shi e Gu, 2021; Zhang *et al.*, 2024), diversos modelos baseados em CNN são aplicados em sensoriamento remoto para diferentes tarefas, como classificação de imagens, detecção de objetos e segmentação semântica. Alguns dos modelos mais utilizados incluem:

- ResNet (ALTAREZ, 2024): Uma arquitetura amplamente utilizada para classificação de imagens e extração de características. No contexto de sensoriamento remoto, tem sido aplicada para tarefas como classificação de uso e cobertura da terra (Dieste *et al.*, 2024), detecção de mudanças (Gong *et al.*, 2022) e mapeamento de áreas urbanas.

- DenseNet (WANG, B. *et al.*, 2024): Uma arquitetura que promove a reutilização de características, melhorando o fluxo de informações na rede. Em sensoriamento remoto, tem sido utilizada para classificação de imagens de alta resolução e detecção de objetos.
- EfficientNet (FERRARI, L. *et al.*, 2021): Uma família de modelos que busca otimizar a acurácia e a eficiência computacional. Em sensoriamento remoto, tem sido aplicada para classificação de imagens de satélite e detecção de alvos específicos, como navios e aeronaves.
- VGG (HAN *et al.*, 2024): Uma arquitetura conhecida por sua simplicidade e bom desempenho em tarefas de classificação de imagens. Em sensoriamento remoto, tem sido utilizada para classificação de uso e cobertura da terra e detecção de mudanças.
- AlexNet (MENG *et al.*, 2024): Uma arquitetura pioneira em Deep Learning, que demonstrou o potencial das CNNs para classificação de imagens em larga escala. Em sensoriamento remoto, tem sido utilizada para detecção de objetos e classificação de cenas.
- U-Net (YANG, Q. *et al.*, 2024): Uma arquitetura popular para segmentação semântica, especialmente em imagens biomédicas. Em sensoriamento remoto, tem sido utilizada para mapeamento de áreas urbanas, detecção de mudanças e extração de características.
- InceptionV3 (ADEGUN; VIRIRI; TAPAMO, 2023a): Uma arquitetura que utiliza múltiplas escalas de convolução para capturar informações contextuais. Em sensoriamento remoto, tem sido aplicada para classificação de imagens de satélite e detecção de objetos.

Mais recentemente, modelos baseados em *transformers* possuem crescente uso (LI, J. *et al.*, 2024), são eles: Swin Transformer (JANNAT; WILLIS, 2022; ZHENG *et al.*, 2023), LAnet (DING; TANG; BRUZZONE, 2021), HRCNet (XU *et al.*, 2020), MANet (MA, X. *et al.*, 2024). Modelos capazes de *self-supervised learning*³ e que também se apresentam por

³ *Self-supervised learning* é um paradigma de aprendizado de máquina onde um modelo é treinado para realizar uma tarefa utilizando os próprios dados para gerar sinais de supervisão, em vez de depender de rótulos fornecidos externamente.

*transfer learning*⁴ também são encontrados na literatura (LI, H. *et al.*, 2021; LI, WENYUAN; CHEN; SHI, 2021).

É importante destacar que a escolha do modelo ideal depende da tarefa em questão, do tipo de dado e dos recursos computacionais disponíveis. Em geral, modelos mais complexos, como o Swin Transformer e o InceptionV3, tendem a apresentar melhor desempenho de maneira geral, mas exigem maior poder de processamento.

No caso específico da mineração, a escolha de modelos mais leves, como a U-Net, pode ser vantajosa, pois permite que o processamento seja realizado em equipamentos mais modestos, democratizando o acesso à tecnologia. Além disso, a U-Net tem se mostrado eficaz em tarefas de segmentação semântica (WANG, RAY; CHOWDHURY; ORTIZ, 2025), o que é crucial para a detecção de áreas de mineração.

Outro fator importante a ser considerado é a disponibilidade de *datasets* de benchmark para mineração. Embora existam *datasets* para outras tarefas em sensoriamento remoto, como classificação de uso e cobertura da terra, a mineração ainda carece de *datasets* abrangentes e com dados rotulados de forma consistente (ADEGUN; VIRIRI; TAPAMO, 2023b). Isso se deve, em parte, à diversidade de respostas espectro-tempo-contextuais e da especificidade espacial das áreas de mineração, o que dificulta a criação de *datasets* genéricos para benchmark.

Nesse contexto, a U-Net se destaca como uma opção promissora para o mapeamento de mineração, pois sua arquitetura flexível e capacidade de aprendizado com poucos dados a tornam adequada para lidar com as particularidades da mineração. Além disso, sua leveza computacional a torna acessível a um público maior, o que contribui para a democratização da tecnologia e o desenvolvimento de soluções mais eficazes para o monitoramento da mineração.

3.3 Seleção de modelos para comparação com o modelo proposto

Esta seção justifica a seleção da arquitetura U-Net como modelo base para a tarefa de segmentação semântica focada na classe de mineração. Argumenta-se em favor da U-Net em

⁴ *Transfer learning*, ou aprendizado por transferência, é uma técnica de aprendizado de máquina onde um modelo treinado para uma tarefa específica é reutilizado como ponto de partida para outra tarefa.

comparação a outras abordagens de Aprendizado Profundo (DL), incluindo arquiteturas mais recentes como Transformers, bem como outros modelos de segmentação explorados na literatura.

Como já foi descrito no início do Capítulo **Erro! Fonte de referência não encontrada.** e na Seção 3.1, a literatura está repleta de modelos capazes de executar a tarefa de segmentação semântica. A escolha entre quais modelos poderiam ser usados é baseada nos seguintes critérios:

- I. **Acessibilidade e Licença de Uso:** O modelo deve ser de código aberto e livremente utilizável. Este requisito é fundamental para garantir a reprodutibilidade integral dos resultados e a aplicabilidade da metodologia desta tese em qualquer território de interesse.
- II. **Eficiência Computacional:** A arquitetura deve possuir reduzida demanda por recursos computacionais. Tal característica visa promover a democratização do método proposto, permitindo sua execução em ambientes com infraestrutura de hardware mais limitada.
- III. **Capacidade Multi-escala e Contextual:** O modelo deve ser capaz de processar e classificar pixels considerando informações em múltiplas escalas e o contexto espacial circundante. Esta capacidade é crucial para o mapeamento em larga escala, como o território nacional, onde a variabilidade no tamanho e distribuição dos alvos (minerações) requer uma abordagem sensível ao contexto local e global, independente de relações fixas baseadas em patches.
- IV. **Paradigma de Aprendizado Supervisionado:** O classificador deve operar no paradigma de aprendizado supervisionado. Esta exigência se alinha com a disponibilidade de dados rotulados pré-existentes, provenientes de bases de dados externas para regiões específicas.

Com base nos critérios estabelecidos e na análise dos subgrupos de arquiteturas para segmentação semântica apresentados na Figura 10, identificou-se que os modelos das subcategorias "Redes Encoder-Decoder" e "Multi-Scale Feature and Pyramid" atendem aos requisitos definidos para este estudo.

A revisão da literatura identificou um número considerável de modelos de DL, tornando inviável a avaliação experimental de todos eles dentro do cronograma deste trabalho de doutorado.

A revisão sistemática da literatura foi realizada consultando dois portais de periódicos principais: Science Direct e Google Scholar. A estratégia de busca empregou a seguinte consulta booleana: (“Segmentation” OR “Detection” OR “mapping” OR “Classification”) AND (“Mining” OR “LAND USE” OR “LAND COVER”) AND (“Brazil” OR “Amazon”). O objetivo foi identificar artigos científicos, publicados entre janeiro de 2023 e fevereiro de 2025, que abordassem a aplicação de técnicas de segmentação semântica, classificação, detecção ou mapeamento com foco nos temas de mineração, uso da terra ou cobertura da terra no Brasil ou, especificamente, na região Amazônica.

A Tabela 3 sumariza os modelos de DL resultantes da revisão da literatura. Eles estão ordenados pela frequência com que foram citados nos artigos encontrados, o que serve para orientar este estudo no reconhecimento das técnicas mais disponíveis e frequentemente empregadas pela comunidade científica. A popularidade de uma técnica não só indica sua democratização e a comprovada capacidade de execução por diferentes autores, mas também sugere sua versatilidade no mapeamento de diversas classes de uso e cobertura da terra.

Tabela 3 – Modelos de Segmentação Semântica encontrados na pesquisa de revisão da literatura.

Nome do Modelo	Contagem de correlatos	Tipo	Subtipo
DeepLabV3+	5.820	Abordagens baseadas em FCN	Rede Encoder-Decoder
DeepLabV3	5.480	Abordagens baseadas em FCN	Convolução dilatada/atrosa
SAM	4.840	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
SegNet	4.690	Abordagens baseadas em FCN	Rede Encoder-Decoder
Unet	4.161	Abordagens baseadas em FCN	Nível de Pixel
ENet	3.340	Abordagens baseadas em FCN	Rede Encoder-Decoder
PSPNet	3.100	Abordagens baseadas em	Recurso

		FCN	multiescala e arquitetura de pirâmide
CANet (Content-Aware Network)	2.310	Abordagens baseadas em FCN	Rede Encoder-Decoder
Segmenter	1.980	Abordagens baseadas em FCN	Rede Encoder-Decoder
Cascade correlation neural network- CCNN	1.170	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
SETR	1.050	Abordagens baseadas em FCN	Convolução dilatada/atrosa
RefineNet	895	Abordagens baseadas em FCN	Rede Encoder-Decoder
ESPNet	853	Abordagens baseadas em FCN	Convolução dilatada/atrosa
SUPM	560	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
One-against-All - OAA	557	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
DeconvNet	315	Abordagens baseadas em FCN	Rede Encoder-Decoder
RANet	285	Abordagens baseadas em proposta regional	Object Detection / Image Labeling
DAFormer	276	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
MCOF	252	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
ParseNet	203	Abordagens baseadas em FCN	Fusão de recursos
SSNet	173	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
LRR-Net	149	Abordagens baseadas em FCN	Recurso multiescala e arquitetura de pirâmide

SMGM	123	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
Context-Aware High-Resolution Domain-Adaptive - HRDA	113	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
SFANet	111	Abordagens baseadas em FCN	Rede Encoder-Decoder
Context Encoding for Semantic Segmentation - EncNet	102	Abordagens baseadas em FCN	Convolução dilatada/atrosa
2D LSTM	85	Abordagens baseadas em FCN	Métodos usando RNN
ExFuse	84	Abordagens baseadas em FCN	Fusão de recursos
Ficklenet	83	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
ReSeg	54	Abordagens baseadas em FCN	Métodos usando RNN
PuzzleCAM	46	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
ECSNet	45	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
STSNet	34	Abordagens baseadas em proposta regional	Object Detection / Image Labeling
Stacked deconvolutional network - SDN	29	Abordagens baseadas em FCN	Rede Encoder-Decoder
GWRP	27	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
DilatedNet	21	Abordagens baseadas em FCN	Convolução dilatada/atrosa
SpyGR	13	Abordagens baseadas em FCN	Recurso multiescala e arquitetura de

			pirâmide
DAG-RNN	13	Abordagens baseadas em FCN	Métodos usando RNN
CGBNet	13	Abordagens baseadas em FCN	Métodos usando RNN
Multiple Description Convolutional - MDC	11	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
DSRG	11	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
CRFasRNN	8	Abordagens baseadas em FCN	Métodos usando RNN
DD-RNN	7	Abordagens baseadas em FCN	Métodos usando RNN
ML-CRNN	4	Abordagens baseadas em FCN	Métodos usando RNN
Convolutional feature masking - CFM	3	Abordagens baseadas em proposta regional	Object Detection / Image Labeling
SA-FFNet	3	Abordagens baseadas em FCN	Fusão de recursos
MIL-FCN	3	Segmentação Semântica Fracamente Supervisionada	Nível de Pixel
CiSS-Net	1	Abordagens baseadas em FCN	Recurso multiescala e arquitetura de pirâmide
MGCRNN	1	Abordagens baseadas em FCN	Métodos usando RNN

De acordo com o ranqueamento da Tabela 3, os seis modelos de DL mais citados, com característica de aprendizado supervisionado e capazes de classificação a nível de pixel serão utilizados para comparação, são eles: DeepLabv3+ (YAO *et al.*, 2025), DeepLabv3 (SUN, YIFEI *et al.*, 2024), SegNet (FU, B. *et al.*, 2021), Enet (GYAWALI *et al.*, 2024; LAO *et al.*, 2024), PSPNet (LI, WEI *et al.*, 2024; SUN, YUCHAO *et al.*, 2025).

A Tabela 4 apresenta os seis modelos de aprendizagem profunda mais citados, ordenados de forma crescente pelo número de parâmetros treináveis em milhões. Essa métrica, que reflete o tamanho e a complexidade do modelo, é comumente utilizada para comparar diferentes arquiteturas de redes neurais. Para fins de democratização de um modelo, quanto menor a quantidade de parâmetros menos recursos computacionais são necessários para sua replicação.

Tabela 4 - Os seis algoritmos mais citados na literatura e os números de parâmetros treináveis em milhões

Modelo	Número de Parâmetros (Milhões)	Topologia
ENet	0,4	Encoder-Decoder. Projetada para eficiência, com um encoder menor e um decoder mais simples. Utiliza convoluções fatoradas e <i>early downsampling</i> para reduzir o custo computacional.
SegNet	28,7	Encoder-Decoder com <i>pooling indices</i>. O encoder extrai características e o decoder as utiliza para gerar a segmentação. Usa <i>pooling indices</i> para preservar informações espaciais durante o <i>upsampling</i> .
Unet	31,06	Formato em U com <i>skip connections</i>. Possui um caminho contrativo (encoder) e um caminho expansivo (decoder) simétricos. As <i>skip connections</i> conectam camadas de resolução similar, ajudando a preservar detalhes finos na segmentação.
DeepLabV3	44,5	Convoluções Atrous (dilatadas) e <i>Atrous Spatial Pyramid Pooling (ASPP)</i>. Usa convoluções atrous para controlar o campo receptivo e capturar informações em múltiplas escalas. O ASPP processa a entrada com diferentes taxas de dilatação, capturando contexto em diferentes tamanhos.
PSPNet	44,5	Pyramid Pooling Module (PPM). Utiliza um módulo de <i>pooling</i> em pirâmide para capturar informações globais e locais em diferentes escalas, melhorando a segmentação de objetos de diferentes tamanhos.
DeepLabV3+	56	Versão aprimorada do DeepLabV3. Adiciona um decoder simples, mas eficaz, para refinar os resultados da segmentação, especialmente nos limites dos objetos. Mantém o ASPP e as convoluções <i>atrous</i> do DeepLabV3.

Buscando o melhor modelo possível para a classificação de mineração em todo Brasil, é feita uma bancada de testes comparativos com os modelos da Tabela 4.

A bancada de testes utilizada no experimento de comparação possui as seguintes especificações: RTX4090 24GB, 128GB DDR4, Ryze 5600 e SSD de 4TB, o experimento usou o mesmo dataset com a mesma subdivisão 70/30 para treino e teste respectivamente. São características desse dataset: 318 mil patches dos quais 21 mil contém a classe de interesse, nesse dataset não foram aplicadas técnicas de *data augmentation*⁵.

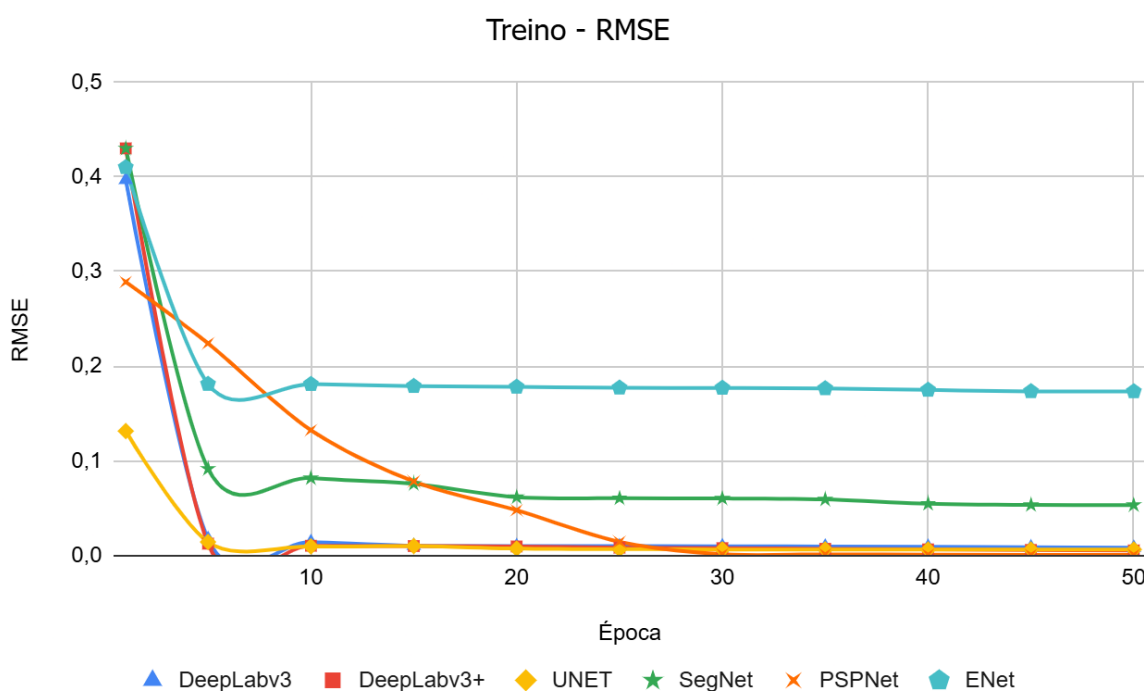


Figura 11 - Comparativo entre algoritmos de DL, treinamento, métrica de Loss por época. Fonte: O Autor (2024)

A Figura 11, apresenta os resultados entre os algoritmos para a métrica de *loss* para cada uma das épocas, o *loss* configurado é o RMSE – *Root Mean Square Error* (HYNDMAN;

⁵ *data augmentation* é uma técnica que aumenta artificialmente a quantidade de dados de treinamento através da aplicação de transformações controladas nos dados existentes, como rotações, espelhamentos e distorções, gerando novas amostras sem coletar novos dados. Isso ajuda a melhorar a capacidade de generalização e a robustez dos modelos de aprendizado de máquina.

KOEHLER, 2006), essa métrica guia o processo de otimização do modelo, que ajusta seus parâmetros internos para minimizar a perda e, conseqüentemente, melhorar a precisão das previsões.

Também é possível inferir pelo gráfico que durante as primeiras épocas é que existe a maior parte da convergência dos algoritmos, exceto a PSPNet em laranja que teve um resultado progressivo em relação ao número de épocas, estabilizando na época 30. Os pontos de mínimo na curva de perda de treinamento mostram as épocas em que o modelo apresentou o menor erro nos dados de treino. Contudo, esses mínimos devem ser interpretados com cautela, pois não necessariamente refletem o melhor desempenho em dados novos. É crucial monitorar a curva de perda de validação em paralelo, pois a divergência entre as duas curvas (treino e validação) é um forte indicativo de *overfitting*. O ponto ideal para interromper o treinamento geralmente ocorre quando a perda de validação começa a aumentar, mesmo que a perda de treinamento continue diminuindo.

A Figura 12 apresenta a comparação dos resultados na etapa de avaliação. Nessa etapa, o algoritmo é utilizado para prever as classes do dataset, e suas previsões são comparadas com os rótulos verdadeiros, já conhecidos por se tratar de um processo de aprendizagem supervisionado. O gráfico apresenta a evolução da métrica de loss durante a fase de validação para seis diferentes arquiteturas de Deep Learning (DL) aplicadas à segmentação semântica. O eixo horizontal representa as épocas de treinamento (de 1 a 50), enquanto o eixo vertical representa o valor do RMSE (de 0 a 0,5000).

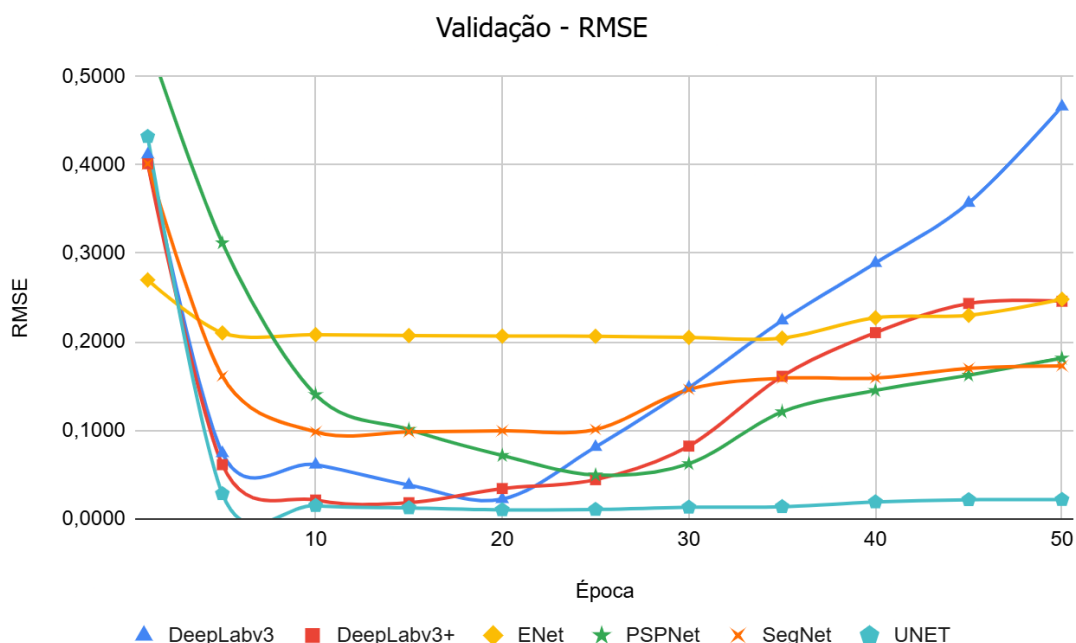


Figura 12 - Comparativo entre algoritmos de DL, validação, métrica de Loss por época. Fonte: O Autor (2024).

Na Figura 12, a U-Net apresenta o melhor desempenho geral, com a menor valor de RMSE na maior parte do gráfico. Começa com valores próxima a 0,45, cai rapidamente para cerca de 0,01 na época 5 e permanece estável, com um leve aumento no final, chegando a cerca de 0,03 na época 50. Mostra que a rede aprendeu eficientemente e manteve um baixo erro de validação.

O modelo DeepLabV3, tem a segunda menor *Loss* na maior parte do treinamento. Inicia com uma *Loss* em torno de 0,45, decresce rapidamente até a época 10, oscila levemente e depois aumenta gradativamente, terminando com uma *Loss* em torno de 0,46 na época 50. Sugere um possível *overfitting* a partir da época 30, onde a *Loss* começa a aumentar.

O Modelo mais atual de DeepLab, a versão plus, mostra um comportamento similar ao DeepLabv3, mas com uma *Loss* ligeiramente maior. Começa próximo a 0,28, cai para menos de 0,1 na época 10, depois sobe gradualmente, chegando a cerca de 0,25 na época 50. Também indica um possível *overfitting* após a época 20.

PSPNet apresenta a terceira menor perda final dentre os modelos avaliados. Embora inicie com uma perda relativamente alta (acima de 0,4), ela converge rapidamente para cerca de 0,05 na época 10. No entanto, a partir desse ponto, observa-se um aumento gradual da perda até aproximadamente 0,15 na época 40, antes de se estabilizar até a época 50. Esse comportamento sugere um aprendizado inicial rápido, seguido por uma possível dificuldade

em refinar ainda mais o ajuste aos dados de treinamento, o que pode indicar uma limitação da arquitetura ou a necessidade de ajustes nos hiperparâmetros. Comparativamente à U-Net, a PSPNet demonstra uma convergência inicial mais lenta, mas atinge um patamar de estabilidade razoável, embora com uma perda final ligeiramente superior.

A ENet em amarelo, mantém uma RMSE relativamente estável ao longo de todo o treinamento, oscilando entre 0,20 e 0,25. Isso sugere que a rede não aprendeu tão bem quanto as outras, ou que atingiu um platô de aprendizado.

Por último, a SegNet: Inicia com uma RMSE em torno de 0,28, cai rapidamente até a época 5, oscila levemente, e depois cresce gradualmente, terminando em cerca de 0,18 na época 50. Apresenta um desempenho intermediário entre o ENet e as demais, também indicando um possível *overfitting*.

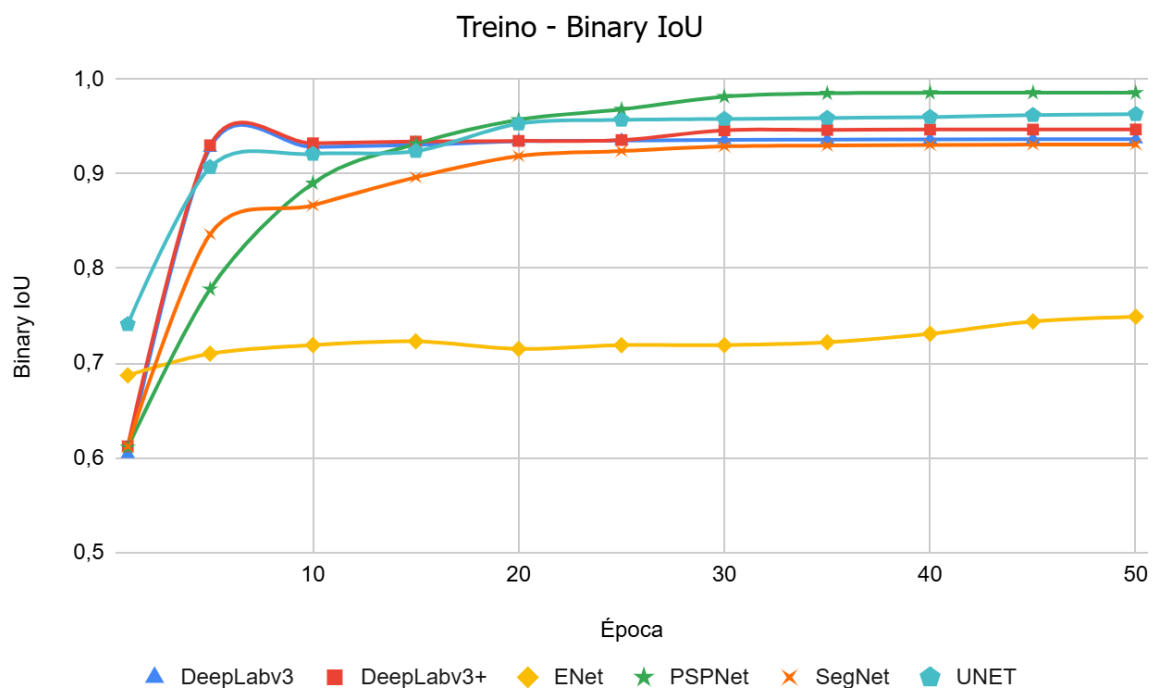


Figura 13 - Comparativo entre algoritmos de DL, treinamento, métrica de *Binary IoU* por época. Fonte: O Autor (2024)

A Figura 13, apresenta a métrica de *Binary IoU* para a etapa de treinamento. O *Binary IoU* (SHELHAMER; LONG; DARRELL, 2017) ou IoU Binário, é uma métrica crucial para avaliar segmentadores semânticos, especialmente quando se trabalha com problemas de classificação

binária, essa métrica mede a sobreposição entre o predito e a verdade. Um valor alto de IoU indica uma boa sobreposição e, portanto, uma segmentação mais precisa e próxima do *ground truth*.

Ainda na Figura 13, o gráfico apresentado compara o desempenho de seis arquiteturas de Deep Learning para segmentação semântica durante a fase de treinamento, utilizando a métrica *Binary IoU*. O eixo horizontal representa as épocas de treinamento (de 1 a 50) e o eixo vertical representa o valor do *Binary IoU* (de 0,5 a 1,0). U-Net e PSPNet se destacam como as arquiteturas com melhor desempenho em termos de aprendizado durante o treinamento, seguidas por DeepLabV3+ e DeepLabv3. O Desempenho do Enet indica que essa arquitetura pode não ser adequada para a tarefa específica ou que precise de modificações em seus hiperparâmetros ou em sua estrutura para que sua aderência ao dataset melhore.

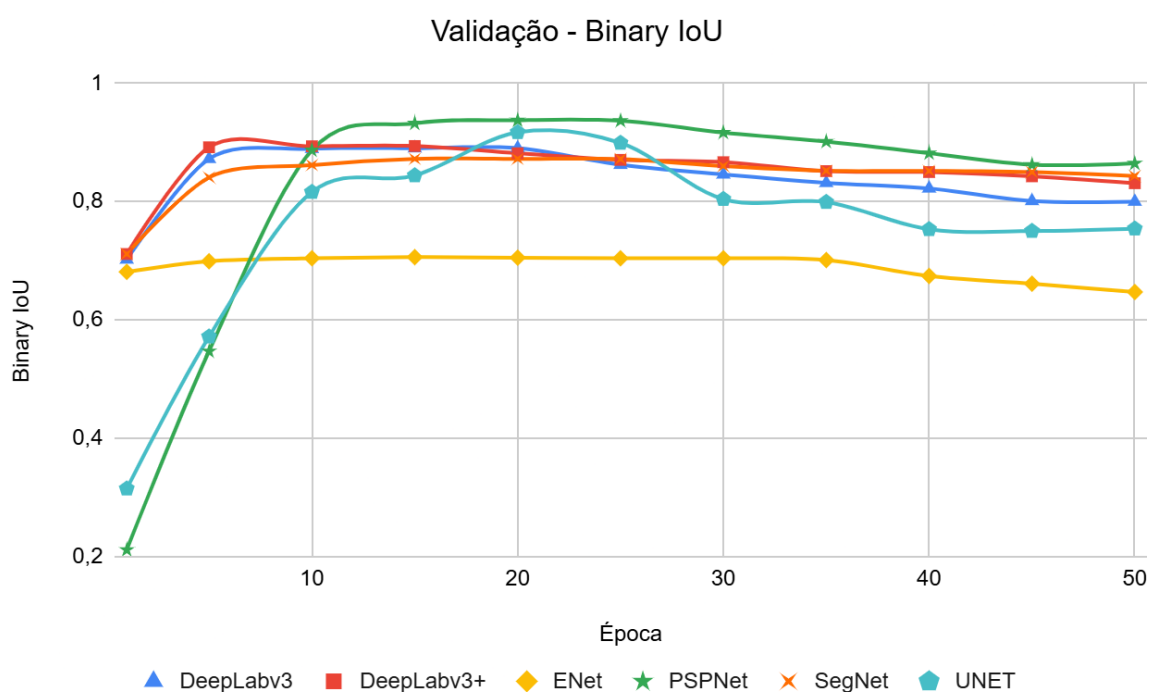


Figura 14 - Comparativo entre algoritmos de DL, validação, métrica de *Binary IoU* por época. Fonte: O Autor (2024).

A Figura 14 avalia a métrica de *Binary IoU* sob perspectiva da etapa de validação, o eixo horizontal representa as épocas de 1 a 50, o eixo vertical apresenta os valores de *Binary IoU* de 0,2 a 1. A PSPNet em verde apresenta o maior valor e, portanto, a melhor aderência

entre o predito e a verdade, o valor sobe rapidamente nas primeiras 15 épocas e apresenta um declínio após a vigésima, o maior valor é de 0,92.

A UNet em azul claro e pentágonos, tem o segundo melhor desempenho. Mostra um crescimento mais lento comparado ao PSPNet, mas consistente. Alcança o melhor IoU de 0,90 na época 20.

O DeepLabV3+ em vermelho e quadrados, apresenta um desempenho próximo ao da U-Net alcançando pico de 0,88 na época 8 e leve tendência de declínio ao longo da série.

O DeepLabV3 em Azul Escuro, mostra um comportamento similar ao DeepLabv3+, mas com um IoU ligeiramente inferior, com máximo em torno de 0,82.

SegNet e Enet tem respectivamente os piores resultados, alcançando máximos abaixo de 0,8, no caso da Enet é notável uma estabilidade próxima a 0,7.

Ao analisar os quatro gráficos em conjunto, o modelo PSPNet demonstra ser a arquitetura que alcança os melhores resultados baseado em seu valor de *Binary IOU* na validação, a U-Net está sempre com valores muito próximos do melhor modelo, com bons valores nos quatro gráficos aqui apresentados, demonstrando estabilidade na segunda posição. Contudo a escolha da melhor arquitetura depende de um balanço entre o desempenho, complexidade e velocidade de treino/inferência, uma vez que no *active learning*⁶ a melhoria do modelo é alcançada através de treinamentos consecutivos com modificações de inclusão e deleção no dataset.

Conforme apresentado na Tabela 4, a arquitetura U-Net demonstrou ser consideravelmente mais eficiente em termos de parâmetros que a PSPNet, possuindo 13.5 milhões de parâmetros treináveis a menos (uma redução de aproximadamente 30% no tamanho total da rede). Apesar dessa diferença significativa no tamanho, a U-Net apresentou uma diferença de apenas 0.05 no *Binary IoU* em comparação com a PSPNet. Diante desse *trade-off* favorável – uma pequena variação no desempenho (IoU) versus um substancial economia de recursos (parâmetros) – a U-Net foi avaliada nesta tese como a arquitetura com o melhor custo-benefício e, portanto, a escolhida para implementação.

⁶ *Active Learning* é uma abordagem de aprendizado de máquina onde o modelo seleciona ativamente os dados mais informativos para serem rotulados por um oráculo (geralmente um humano), visando melhorar seu desempenho com menos dados rotulados. Em resumo, o modelo "aprende a perguntar" quais dados são mais úteis para seu aprendizado.

4 Método Proposto

O presente capítulo descreve a metodologia desenvolvida para atingir os objetivos desta pesquisa, concentrando-se na classificação da mineração em larga escala espacial e temporal no Brasil por meio de técnicas de DL. A Seção 4.1 detalha as modificações realizadas na arquitetura U-Net (RONNEBERGER; FISCHER; BROX, 2015) escolhida na seção 3.3. Tais modificações foram projetadas para adaptar a U-Net às particularidades dos dados de sensoriamento remoto multitemporais e para aumentar a eficiência computacional, visando facilitar a aplicação em larga escala e a reprodutibilidade.

Posteriormente, a Seção 4.2 introduz a proposição central deste trabalho: um método sistemático e automatizado para o mapeamento da mineração em séries temporais. Este método híbrido combina o modelo de classificação adaptado com as capacidades de processamento em nuvem do *Google Earth Engine* e a otimização de dados via formato TFRecord, buscando a eficiência computacional para leitura e escrita de todo dataset e permitir a análise de vastos conjuntos de dados geoespaciais multitemporais.

4.1 Modificações na U-Net

Embora a U-Net original tenha sido identificada como a arquitetura de melhor custo-benefício para esta pesquisa (Seção 3.3), sua aplicação direta em larga escala, considerando o processamento de uma série temporal de 37 anos para todo o território nacional com dados multiespectrais e multitemporais, apresenta desafios computacionais significativos. Tais desafios tornam os processos de treinamento e classificação dependentes de recursos de hardware consideráveis, incluindo GPUs dedicadas com pelo menos 12GB de VRAM, o que pode limitar a replicabilidade da metodologia proposta em ambientes com infraestrutura computacional mais modesta.

Esta demanda por recursos computacionais é inerente a arquiteturas convolucionais como a U-Net ao processar dados com alta dimensionalidade de entrada. O número de parâmetros e o custo computacional (FLOPs - *Floating Point Operations* ou Operações em Ponto Flutuante) de uma camada convolucional, especialmente a primeira, escalam linearmente com o número de canais de entrada. Por exemplo, processar dados com 6 canais de entrada em vez de 3

aproximadamente dobra os parâmetros e FLOPs associados à primeira camada convolucional e contribui significativamente para o custo computacional geral do modelo, impactando diretamente os requisitos de memória (VRAM).

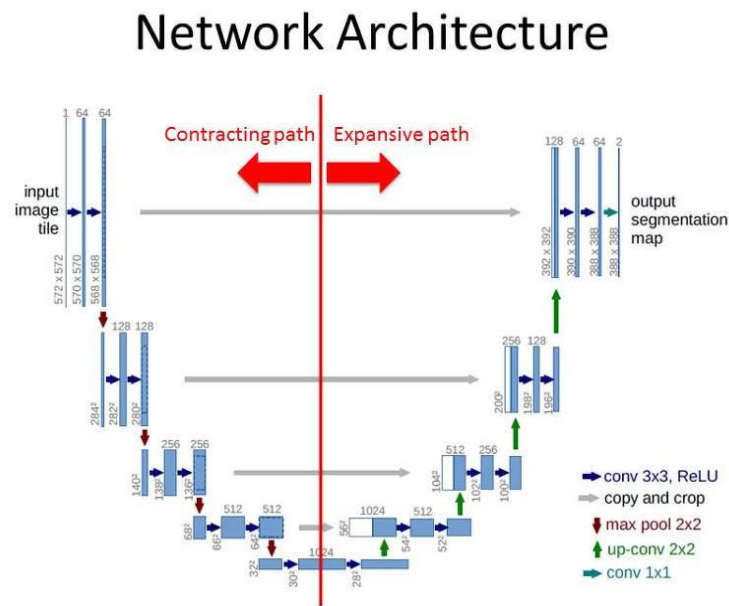


Figura 15 - Arquitetura original da U-Net. Fonte: (RONNEBERGER; FISCHER; BROX, 2015).

Visando mitigar o esforço computacional e as demandas de VRAM da U-Net original, cuja arquitetura é ilustrada na Figura 15, para aplicação em larga escala, propõe-se uma versão modificada da arquitetura, apresentada na Figura 16. As estratégias adotadas buscaram reduzir a complexidade e a dimensionalidade da rede, com o objetivo de simplificar o algoritmo e otimizar seu desempenho sem comprometer significativamente a acurácia. As principais modificações implementadas foram:

- I. Aplicação de Padding por Espelhamento: Para remover o efeito de perda de pixels nas bordas das imagens após as operações convolucionais e manter o tamanho espacial uniforme em cada nível da arquitetura em "U". Esta técnica, conhecida como *reflection padding* ou espelhamento matricial nas bordas, evita a perda de informação nas extremidades dos *chips* de entrada e elimina a necessidade de operações de recorte ou convoluções adicionais para ajuste dimensional.
- II. Redução da Profundidade da Arquitetura: A profundidade da rede, referente ao número de pares de camadas de *downsampling* e *upsampling*, foi diminuída. Esta

modificação foi validada experimentalmente no domínio da aplicação para identificar a configuração que apresentasse a menor perda de acurácia em comparação com a estrutura padrão da U-Net, buscando um equilíbrio entre eficiência e desempenho.

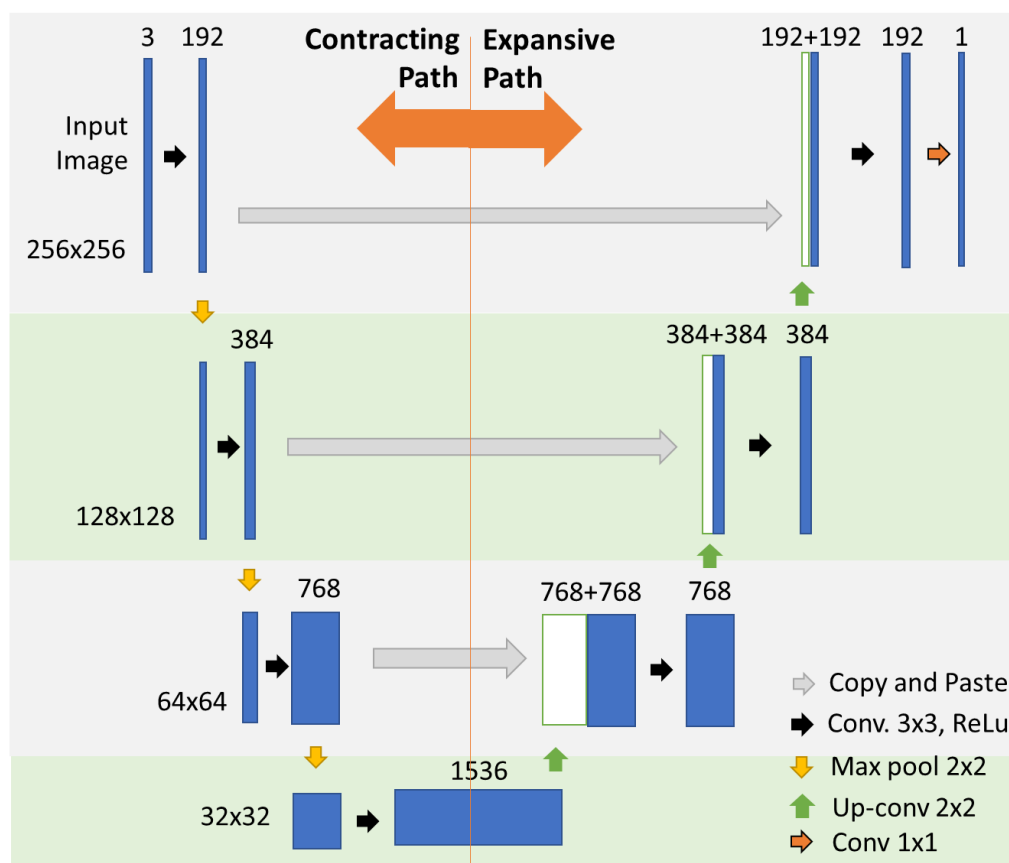


Figura 16 - Estrutura da “U-Net modificada” utilizada para classificação neste trabalho. Fonte: O Autor (2024).

Para avaliar as modificações implementadas na arquitetura da U-Net, um procedimento experimental específico foi desenvolvido e é ilustrado na Figura 17. Este procedimento consistiu em:

- I. Seleção Aleatória de Regiões: Selecionar aleatoriamente um conjunto de 10 regiões de 10 km x 10 km. Essa seleção visou garantir que os dados utilizados para a validação das modificações fossem representativos.
- II. Dataset de Treinamento Único: Construir um único dataset de chips a partir dessas regiões, utilizando-o para treinar tanto a U-Net original quanto a versão

modificada. O uso do mesmo conjunto de dados é fundamental para assegurar uma comparação equitativa do desempenho das duas arquiteturas.

- III. Classificação da Série Temporal: Aplicar os modelos treinados aos dados correspondentes aos 37 anos da série temporal para as regiões selecionadas.
- IV. Comparativo de Área Preditada: Comparar o desempenho das duas versões da U-Net calculando a área total predita para a classe de mineração anualmente ao longo do período. Este comparativo de área permite verificar se a U-Net modificada produz resultados de classificação espacialmente consistentes e quantitativamente similares aos da U-Net original ao longo do tempo, apesar de suas diferenças arquiteturais e redução de complexidade.

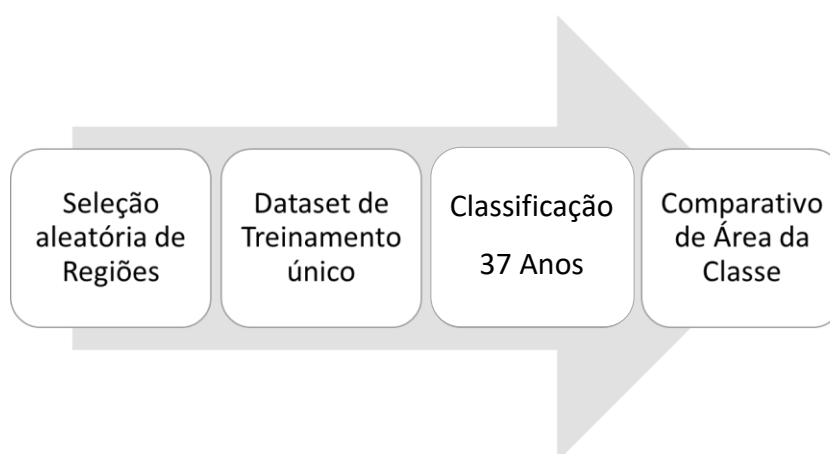


Figura 17 - Procedimento para validação da arquitetura modificada. Fonte: O Autor (2024).

Os resultados do procedimento de validação das modificações arquiteturais são apresentados na Figura 18, que compara a variação percentual da área predita para a classe de mineração entre a U-Net modificada e a U-Net original ao longo dos 37 anos da série temporal. A análise do gráfico revela que a maior parte dos pontos de diferença percentual se concentra próxima a zero, indicando que a área predita pela versão modificada é consistentemente similar à área predita pela U-Net original na maioria dos anos. Observam-se oscilações nos valores de diferença, alternando entre previsões ligeiramente maiores ou menores, mas sem um padrão consistente de superestimação ou subestimação ao longo do tempo. A linha de tendência, apesar de levemente ascendente, exibe uma inclinação quase horizontal, com um baixo valor de R^2 (0,079). Este resultado reforça que a variação entre os modelos é predominantemente

aleatória e não indica um desvio sistemático significativo no que tange à predição de área ao longo da série temporal.

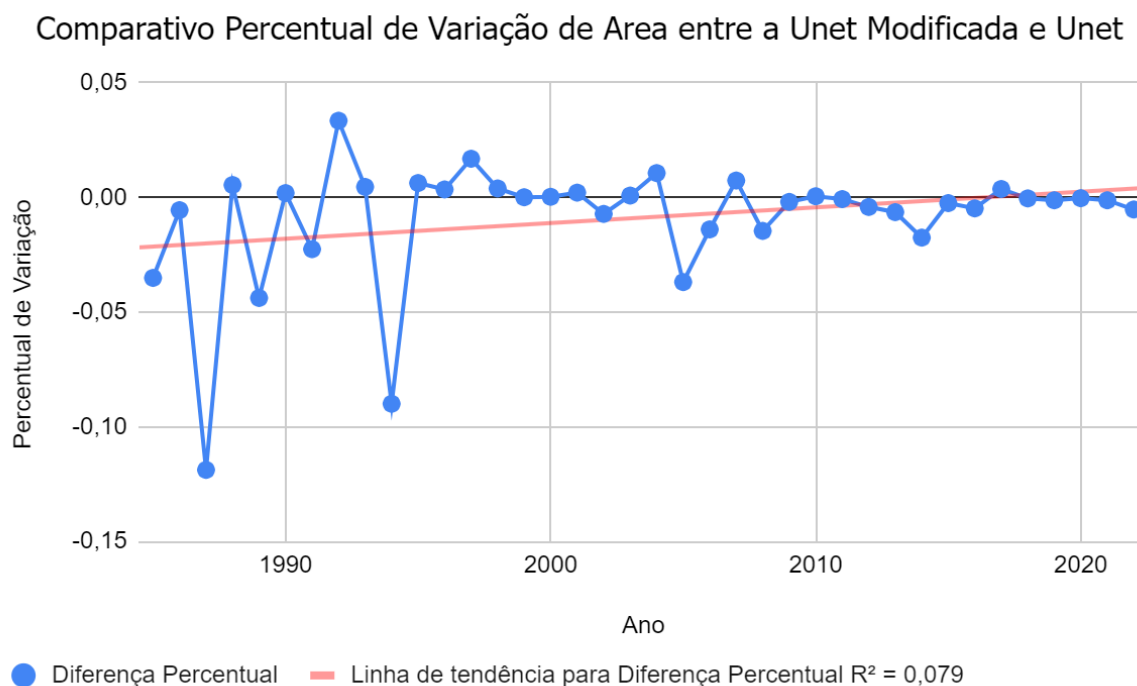


Figura 18 - Gráfico comparativo de variação percentual de área entre o predito anual da U-Net modificada e a U-Net original de (RONNEBERGER; FISCHER; BROX, 2015). Fonte: O Autor (2024).

Desta forma, demonstra-se que, apesar da redução de aproximadamente 30% no número de parâmetros, a U-Net modificada possui habilidades de aprendizado e capacidade de predição de área muito próximas às da U-Net original para a tarefa de classificação binária em mosaicos anuais Landsat livres de nuvem, conforme utilizado nesta tese. Esta similaridade de desempenho, aliada à sua maior eficiência computacional, valida as modificações implementadas e a torna adequada como arquitetura base para o método proposto. A seguir, a Seção 4.2 detalha como esta arquitetura adaptada é integrada em uma metodologia híbrida completa para o mapeamento em larga escala da mineração no Brasil.

4.2 Arquitetura Híbrida para Classificação Temporal e em Larga Escala da Mineração no Brasil

Esta seção introduz a proposição central deste trabalho: um método híbrido, sistemático e automatizado para o mapeamento da mineração em séries temporais em larga escala espacial e temporal. A metodologia proposta, ilustrada em seu *pipeline* completo na Figura 19, está estruturada em quatro áreas principais, cada uma englobando subtarefas específicas: Geração de Mosaicos, Dataset de Treinamento, Classificação e Integração, e Acurácia.

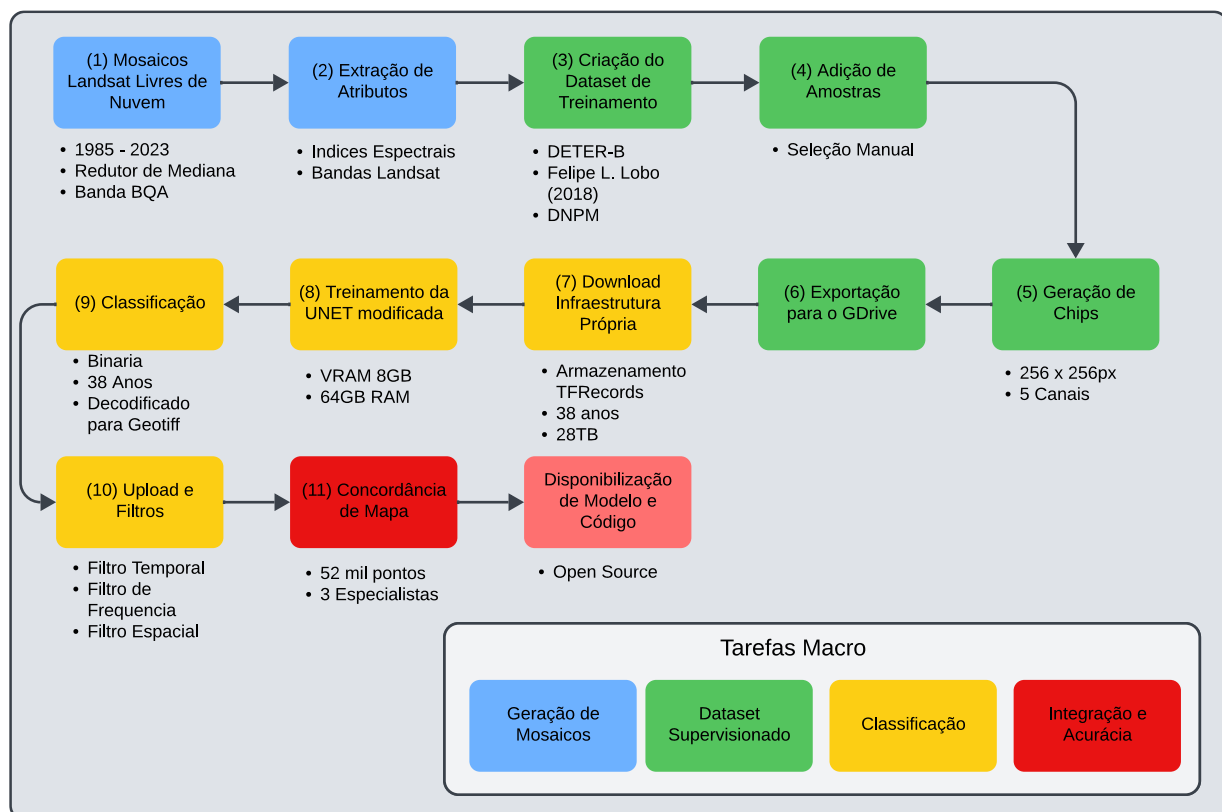


Figura 19 - Metodologia Aplicada para o Mapeamento das Minerações no Brasil. Fonte: O Autor (2024).

A primeira área, Geração de Mosaicos, compreende a criação de mosaicos anuais livres de nuvens a partir das imagens de satélite Landsat para o período de 1985 a 2022. Este processo inicial envolve a utilização da banda de qualidade de pixel (*Band Quality Assessment* - BQA) para a identificação e mascaramento de nuvens e sombras, e a aplicação de um redutor de

mediana à pilha de imagens de cada ano para gerar um mosaico representativo e livre de ruídos atmosféricos. Adicionalmente, nesta etapa, são calculados e incluídos nos mosaicos índices espectrais como o *Normalized Difference Vegetation Index* - NDVI (YANG, Y. *et al.*, 2018), para realçar a vegetação, o *Normalized Difference Soil Index* - NDSI (MADASA; ORIMOLOYE; OLOLADE, 2021) para evidenciar o solo, e o *Normalized Difference Water Index* - NDWI (GAO, B. C., 1996) para identificar corpos d'água. Embora modelos de DL possuam a capacidade de extração automática de características, a inclusão destes índices pré-calculados no *input* visa fornecer informações adicionais relevantes desde o início, potencialmente facilitando o aprendizado e acelerando a convergência da rede.

A segunda área da metodologia, Dataset de Treinamento, concentra-se na criação eficiente do conjunto de dados supervisionado, minimizando a necessidade de rotulagem manual extensiva. Para tanto, foram utilizadas fontes de dados geoespaciais existentes como base inicial: o mapeamento de minerações por (LOBO, F. DE L. *et al.*, 2018) para a Amazônia em 2018, alertas de mineração do projeto DETER (DINIZ, C. G. *et al.*, 2015), e vetores de mineração do SIGMINE/DNPM.

O processo de amostragem foi realizado dentro dos vetores de “Autorização de Lavra” definida pelo shapefile do SIGMINE/DNPM e dos dois outros trabalhos fontes, expandida por um *buffer* de 10km. Dentro dessas áreas delimitadas, aplicou-se o algoritmo K-means com $k=4$ para clusterização dos pixels nos dados de 2018 na plataforma Google Earth Engine (GEE). Os clusters resultantes que melhor representaram as áreas de mineração foram subsequentemente selecionados e binarizados, gerando uma classificação binária preliminar para o ano de 2018, distinguindo mineração de outras classes.

Esta classificação binária preliminar, obtida de forma semi-automatizada, serviu como rótulo base para a amostragem final do dataset supervisionado. Chips quadrados de 256 pixels foram amostrados dentro das áreas classificadas, constituindo o conjunto de dados de treinamento para a U-Net modificada. O dataset resultante, abrangendo amostras para todo o território brasileiro, totaliza 1,90TB e foi exportado no formato TFRecord para otimizar o armazenamento e a leitura eficiente pela rede durante a etapa de treinamento subsequente.

A terceira área da metodologia, Classificação e Integração, conforme ilustrado na Figura 19, abrange o treinamento do modelo de classificação e a aplicação deste para gerar os mapas de mineração anuais. As subtarefas iniciais (Subtarefa 8) envolvem o download do dataset de

treinamento (1.90TB em formato TFRecord, como descrito anteriormente) e de todos os 37 mosaicos anuais livres de nuvens, totalizando aproximadamente 22TB. A leitura eficiente desses 22TB de dados é um ponto crítico superado pelo uso do formato TFRecord, que otimiza a leitura em disco pelo framework Tensorflow permitindo o paralelismo no acesso durante os processos subsequentes.

O treinamento da U-Net modificada (Subtarefa 9) foi executado utilizando dois *scripts* principais, um para treino e outro para classificação. Para o treinamento, foi utilizado um servidor próprio com a seguinte configuração ("máquina de produção"): Linux Debian, 16 núcleos de processamento, 128 GB de RAM a 3200Mhz, 40 TB de HD, 12TB de SSD e uma placa gráfica Nvidia RTX 4090 com 24GB de VRAM. Adicionalmente, testes foram conduzidos em um ambiente mais simples para verificar os requisitos mínimos: Linux Debian, 4 núcleos, 64GB de RAM, 1TB de SSD e uma RTX 2070 Super (8GB). O treinamento na máquina de produção teve duração de 18 horas, finalizando na época 130. O critério de parada para o treinamento foi definido pelo decaimento mínimo de 0,001 no índice Binary IoU na validação. Para evitar o *overfitting*, adotou-se a técnica de *early stopping* (PRECHELT, 2012), selecionando o modelo da época 110 por apresentar o melhor equilíbrio entre desempenho e generalização.

Antes da classificação em larga escala, uma etapa de validação dos resultados do treinamento foi realizada por uma equipe de três especialistas em Geografia, Oceanografia e Geologia, treinados no reconhecimento da classe alvo. Após a avaliação e validação dos resultados preliminares pelos especialistas, iniciou-se o processo de classificação dos 37 anos de dados em larga escala (Subtarefa 10).

Na máquina de produção, a classificação de cada ano da série temporal levou aproximadamente 38 minutos. Em seguida, foram necessários 20 minutos adicionais para a decodificação do *output* do modelo (inicialmente em formato TFRecord) para o formato *raster* GeoTIFF. Este dado em GeoTIFF serve como *input* para a próxima etapa do pipeline (Subtarefa 11), que envolve o upload dos resultados para o Google Earth Engine (GEE). O processamento completo do dataset (classificação dos 37 anos) e o upload de todos os resultados anuais para o GEE totalizaram aproximadamente 34 horas.

Uma vez que toda a série temporal é classificada, os resultados brutos passam por uma etapa crucial de pós-processamento e filtragem. Esta etapa é realizada diretamente na plataforma Google Earth Engine (GEE) e tem como objetivo primordial garantir a consistência espacial e temporal dos mapas anuais de mineração gerados. O processo de filtragem visa eliminar *outliers* (classificações pontuais inconsistentes) e corrigir transições temporais que não são representativas da dinâmica real da atividade mineradora. Para tal refinamento, são aplicados três filtros distintos em sequência.

O primeiro é o Filtro Espacial, cuja responsabilidade é remover agrupamentos de pixels classificados como mineração cuja área total é inferior a 10 unidades espaciais (por exemplo, 10 pixels de 30x30m, totalizando 9000 m²). Este filtro baseia-se na premissa de que áreas de mineração reais, mesmo as de menor escala, tendem a apresentar uma extensão espacial mínima coerente, permitindo assim eliminar classificações incorretas esparsas ou ruídos isolados espacialmente.

O segundo é o Filtro Temporal, que visa garantir a congruência temporal da classe mapeada. Este filtro opera analisando a classificação de cada localização em uma janela temporal móvel (neste caso, avaliando sempre em intervalos de 3 anos consecutivos), eliminando classificações que apareçam ou desapareçam de forma abrupta e sem persistência mínima nesse período.

O terceiro e último é o Filtro de Frequência, atua na remoção de transições rápidas ou ruídos de classificação que não se mantêm temporalmente por mais de 3 anos em uma mesma localização espacial, reforçando a estabilidade e a persistência da classe de mineração mapeada ao longo da série temporal.

Na sequência do processo de filtragem, as áreas classificadas como mineração passam por uma etapa de categorização visual. Uma equipe de três especialistas, com conhecimento na identificação de padrões de mineração, avalia visualmente as áreas mapeadas e as categoriza entre os subtipos 'Garimpo' e 'Indústria'.

Com o dado classificado, filtrado e categorizado visualmente, ele está pronto para ser comparado com dados de referência preexistentes para fins de validação e análise. Essas comparações incluem o confronto dos resultados com os dados do DNPM para o ano de 2010

e com o mapeamento de Lobo et al. (2018) para 2018. Esses resultados refinados e validados servem como base para as análises apresentadas no Capítulo 5.

5 Resultados

Os resultados aqui apresentados são visualizações e métrica dos outputs do método apresentado nessa tese, onde todo o território nacional pôde ser classificado, com a sua acurácia calculada por grids de 100 x 100 km, conforme apresentado na seção 5.3, o comprometimento desde classificador é que ele possua um bom desempenho para todo o território nacional, em especial para áreas que tenham presença de atividade garimpeira com indícios de ilegalidade, conforme o descrito na seção 5.2.

5.1 Evolução da área de mineração: Mineração industrial e Garimpo

A extensão, em superfície, da mineração brasileira foi monitorada sistemática e continuamente, produzindo mapas e estatísticas anuais de 1985 a 2022. Mais ainda, os dados produzidos permitem a classificação de dois padrões distintos de extração: o garimpo e a mineração industrial Figura 20. Os dados aqui apresentados fornecem uma melhor compreensão espacial da dinâmica da extração de mineração no Brasil, abrindo caminho para uma maior compreensão de seus impactos socioambientais e atualizando o status dessa atividade econômica por meio de 37 anos de dados.

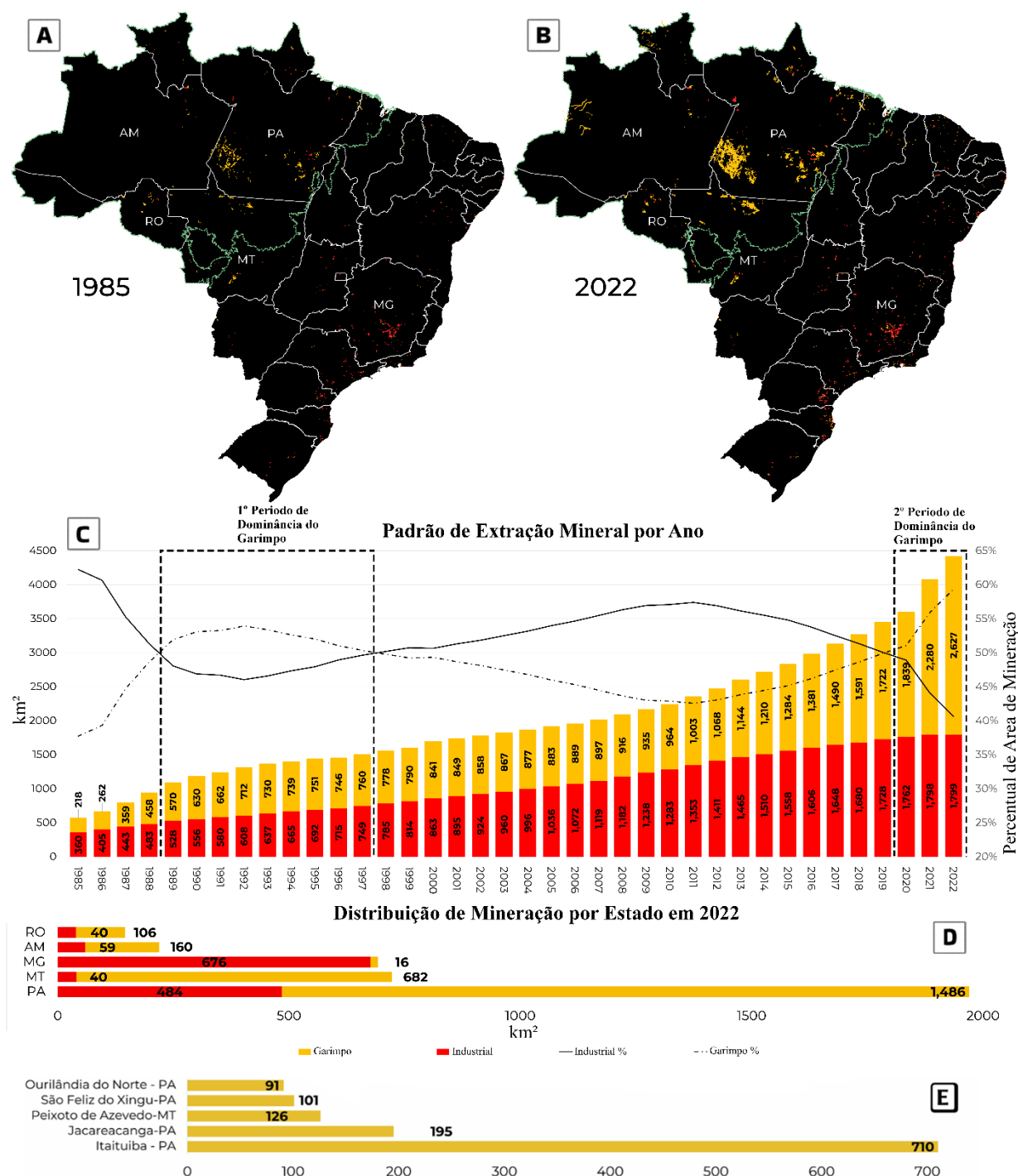


Figura 20 - Distribuição espacial da mineração no Brasil, apresentado em 5 gráficos: A) Distribuição espacial em 1985; B) Distribuição espacial em 2022; C) Área total da mineração por ano e por tipo; D) Os 5 estados com maior área de mineração em 2022; E) Os 5 municípios com maior área de mineração de garimpo em 2022. Fonte: O Autor (2025).

Em termos de padrões de extração, o Brasil exibiu um perfil de exploração majoritariamente industrial por quase quatro décadas. No entanto, o padrão recente mudou, e

a área ocupada por atividades de garimpo em 2022 é 1,46 vezes maior do que a encontrada para extração industrial. Ao longo de quase quatro décadas, o Brasil exibiu dois períodos distintos em que a área de extração garimpeira excedeu a área de extração industrial, de 1989 a 1997 e de 2020 a 2022 Figura 20 gráfico C.

No início da série temporal, em 1985, a área ocupada por operações de mineração industrial atingiu aproximadamente 360 km², enquanto a área ocupada por operações de extração de garimpo atingiu aproximadamente 220 km². Além da reversão do domínio da mineração industrial de 2020 para 2022, houve uma aceleração notável na expansão do garimpo. Tal extração aumentou notavelmente no espaço, dominando o norte do Brasil, principalmente no Pará, Mato Grosso, Rondônia e Amazonas. Além disso, a atividade industrial concentrou-se na região sudeste do Brasil, especialmente em Minas Gerais, Figura 20 gráfico D.

Em 2022, a área de extração de garimpo atingiu 2630 km², 12 vezes maior do que em 1985, um aumento de aproximadamente 1200%. Além do aumento, em 2022, 91,75% (2410 km²) da área de garimpo está dentro do limite do bioma Amazônia Figura 20 gráfico E. É importante ressaltar que a área associada à mineração industrial também aumentou no Brasil, mas de forma mais gradual e em uma intensidade muito menor do que nos anos mais recentes da série temporal relativa à expansão garimpeira. Em 1985, a área de mineração industrial atingiu 360 km², mas aumentou para 1800 km² em 2022, representando um aumento de 5 vezes.

No entanto, é a taxa de expansão dos últimos anos o aspecto mais alarmante do crescimento garimpeiro. Nos últimos quatro anos da série, ou seja, de 2019 a 2022, quase não houve expansão relevante da área de mineração industrial, que aumentou de 1.730 km² em 2019 para 1.800 km² em 2022, um aumento absoluto de aproximadamente 4%. Em contraste, se a estabilidade é uma marca registrada recente da extração industrial, o exato oposto se aplica à área de extração garimpeira. De 2019 a 2022, a área de extração garimpeira aumentou de 1.720 km² em 2018 para 2.670 km² em 2022 — um aumento absoluto de 55%. Além disso, no último ano da série, ou seja, de 2021 para 2022, a área garimpada aumentou em 350 km², passando de 2.280 para 2.630 km².

O estado do Pará é de relevância central para a expansão observada na Amazônia. O estado contém quatro dos cinco maiores municípios do Brasil em termos de extensão da área garimpada. Notavelmente, Itaituba (710 km²), Jacareacanga (195 km²), São Feliz do Xingu (101 km²) e Ourilândia do Norte (91 km²) constituem conjuntamente 41,7% da área nacional

de garimpo, com uma área de 1097 km². Em terceiro lugar, o único município fora do Pará é Peixoto de Azevedo, no Mato Grosso (MT), com 126 km² de atividades de garimpo em 2022.

5.2 Indícios de Ilegalidade e Persistência Temporal de Sítios Garimpeiros

Recentemente, a atividade de garimpo se espalhou por toda a Amazônia. Inicialmente concentrada na região sul do Pará e nas partes norte dos estados de Mato Grosso e Rondônia Figura 20 gráfico B, a extração de garimpo atingiu a fronteira de países vizinhos (Guiana Francesa, Venezuela e Colômbia), o interior de UCs de proteção integral, UCs de uso sustentável, reservas do patrimônio natural (RPPNs), reservas extrativistas (RESEXs) e TIs, conforme mostrado na Figura 21. Em 2022, 780 km² de garimpos ocorreram dentro de UCs, enquanto 250 km² ocorreram dentro de TIs, embora sua distribuição, em ambos os casos, não tenha sido homogênea.

Quando ranqueados por área, cinco UCs foram responsáveis por 85,5% da área total garimpada no interior de UCs: APA do Tapajós (área de proteção ambiental - APA), FLONA do Amana (floresta nacional - FLONA), ESEC Juami-Japura (estação ecológica - ESEC), FLONA do Crepori e PARNA do Rio Novo (parque nacional - PARNA). Da mesma forma, os sinais de atividade garimpeira não foram distribuídos uniformemente entre as TIs do Brasil. As TIs Kayapó, Munduruku e Yanomami foram responsáveis por mais de 90% da área total de extração garimpeira no interior de territórios indígenas Figura 21.

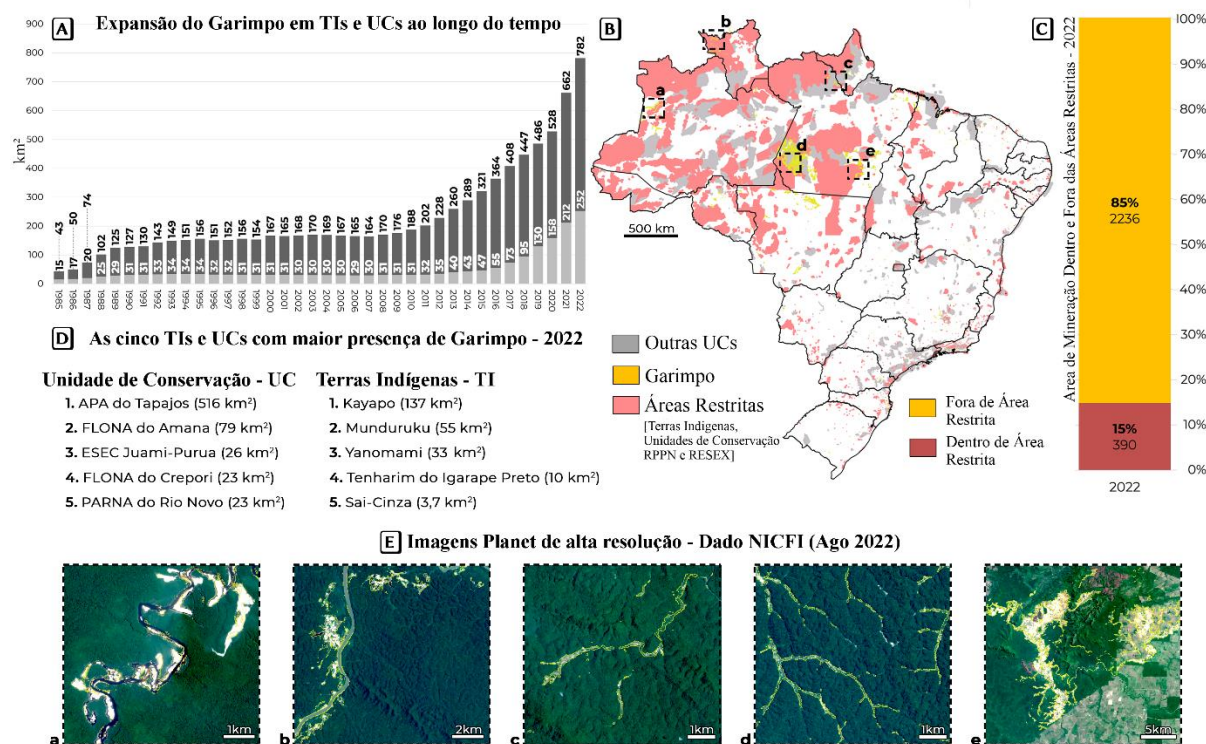


Figura 21 - Garimpos em áreas restritas, 3 gráficos: A) Serie temporal da expansão do garimpo em áreas protegidas; B) Mapa com os garimpos em áreas protegidas no Brasil; C) 5 Terras indígenas e 5 Unidades de Conservação com maior presença de garimpo em 2022; e D) Extensão do garimpo dentro e fora das áreas protegidas. Fonte: O Autor (2024).

Em 2022, pelo menos 15% (390 km²) da área de garimpo, como mostrado na Figura 21 gráfico D, ocorreu dentro de regiões com restrições estabelecidas para limitar essa atividade. A atividade de extração mineral, de qualquer tipo, é constitucionalmente proibida dentro de TIs, UCs de proteção integral, RPPNs e RESEXs. Além disso, proporcionalmente à área detectada na última década (2013-2022), a expansão do garimpo dentro de TIs e UCs foi 462% maior em 2022 do que em 2013. Somente dentro de TIs, a área de garimpo em 2022 foi 625% maior do que em 2013, um ganho absoluto de 210 km², saltando de 40 para 250 km². Dentro das UCs, a área de garimpo foi 300% maior em 2022 do que em 2013, com um ganho absoluto de 520 km², saltando de 260 para 780 km². Combinada, a expansão da mineração em TIs e UCs atingiu 730 km² em uma década.

O cruzamento de locais de garimpo com as bases nacionais de regiões com restrição de mineração é uma maneira simples de verificar a ilegalidade espacial, conforme Lei nº 11.685/2008 62. No entanto, violações espaciais não são o sinal mais comum de ilegalidade. Uma operação de mineração pode apresentar sinais de ilegalidade mesmo quando não está localizada em uma região restrita. Uma maneira de verificar isso é usando a plataforma

SIGMINE. Desenvolvida e atualizada pela Agência Nacional de Mineração (ANM), esta plataforma fornece a localização e os detalhes dos processos minerários ativos em todo o Brasil.

A plataforma SIGMINE adiciona um grau essencial de transparência às informações da ANM, permitindo a verificação pública de registros formais de extração mineral no Brasil. Vale ressaltar porém, que mesmo quando registrado na plataforma SIGMINE como titular formal de uma permissão de garimpo ou de uma concessão de lavra, a ausência de uma licença ambiental preexistente, emitida por um instituto ambiental brasileiro oficial (Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis — IBAMA — ou a Secretaria Estadual do Meio Ambiente — SEMA — e no caso do estado do Pará, a Secretaria Municipal do Meio Ambiente — SEMA) também pode categorizar os operadores como envolvidos em atividades ilegais. Nesse sentido, todos os locais de garimpo detectados foram classificados como apresentando sinais explícitos de ilegalidade ou sem sinais explícitos de ilegalidade, confiando exclusivamente na posição do local de extração de garimpo versus a geolocalização e o tipo de licença registrados na plataforma SIGMINE. Uma verificação mais aprofundada do licenciamento ambiental, garantindo a preexistência de uma autorização ambiental para cada lote registrado no SIGMINE, não pôde ser conduzida, pois nenhum banco de dados centralizado ou digital permite acesso público a todas as licenças ambientais emitidas para atividades de mineração.

5.3 Análise de Acurácia do Mapa

A avaliação da acurácia de um mapa classificado emerge como um procedimento essencial para determinar a fidedignidade e a aplicabilidade do produto cartográfico final. Embora a acurácia do classificador se concentre na proficiência do algoritmo em si para rotular corretamente pixels individuais, a acurácia do mapa classificado avalia o resultado abrangente do processo de mapeamento. Em outras palavras, busca-se quantificar o grau de correspondência entre o mapa e a distribuição real das classes de interesse no terreno.

Essa distinção se revela fundamental, visto que um classificador pode exibir um desempenho satisfatório em testes com dados de treinamento e de avaliação, contudo, o mapa gerado pode conter imprecisões em decorrência de fatores como ruído nos dados, complexidade da paisagem ou limitações na resolução espacial das imagens.

A mineração, apesar de sua importância histórica e relevância econômica no Brasil, é considerada um padrão de uso da terra estatisticamente raro, dada a vasta área geográfica do país (MACHADO; SILVIA, 2001). Essa característica impõe desafios significativos para o projeto de amostragem, tanto no treinamento de modelos quanto na avaliação de erros. De acordo com Olofsson et al., (2014), a abordagem de amostragem aleatória simples, estratificada apenas por região, poderia resultar em erros de omissão substanciais, comprometendo a precisão da avaliação gerada. Esses erros ocorrem quando áreas de mineração reais não são identificadas corretamente no mapa, levando a uma subestimação da extensão da atividade.

Seguindo o guia de boas práticas proposto por Olofsson et al., (2014), para mitigar o impacto desses erros de omissão e comissão, foi adotada a abordagem de amostragem aleatória estratificada em dois estágios.

Inicialmente, uma malha de grades foi sobreposta ao território brasileiro, dividindo-o em células de 250 km x 250 km, totalizando 182 células. A existência dessa grade, por si só, justifica a necessidade de uma distribuição aleatória de um número específico de amostras por célula, mesmo naquelas onde não há indícios de pixels de mineração. Essa primeira etapa garante uma cobertura amostral aleatória abrangente do território.

A segunda etapa da amostragem estratificada em dois estágios considera as proporções de área dos estratos de mineração e não mineração. A área de cada classe, nesse contexto, atua como um fator de ponderação, influenciando a distribuição variável de amostras dentro de cada uma das duas classes identificadas: mineração e não mineração. Essa ponderação garante que áreas com maior proporção de mineração recebam um número de amostras proporcionalmente maior, otimizando a detecção de áreas mineradas e, conseqüentemente, reduzindo a probabilidade de perder erros de omissão.

A definição do número de amostras por grade e por classe, detalhada na Figura 22, foi realizada com base na Equação 1, proposta por Olofsson et al. (2014):

Equação 1 – Cálculo da quantidade de amostras por grid e por classe, obtida em (OLOFSSON *et al.*, 2014)

$$N = \left(\frac{\sum_i W_i S_i}{S(\hat{\theta})} \right)^2$$

Nessa equação, $S(\hat{\theta})$ representa o desvio padrão esperado na acurácia global, adotado como 0,5%, enquanto W_i é a proporção da área mapeada da classe i , e S_i corresponde ao desvio

padrão esperado da classe i. A aplicação dessa fórmula para cada uma das 182 células resulta em um total de 52.320 amostras, distribuídas por todo o território brasileiro.

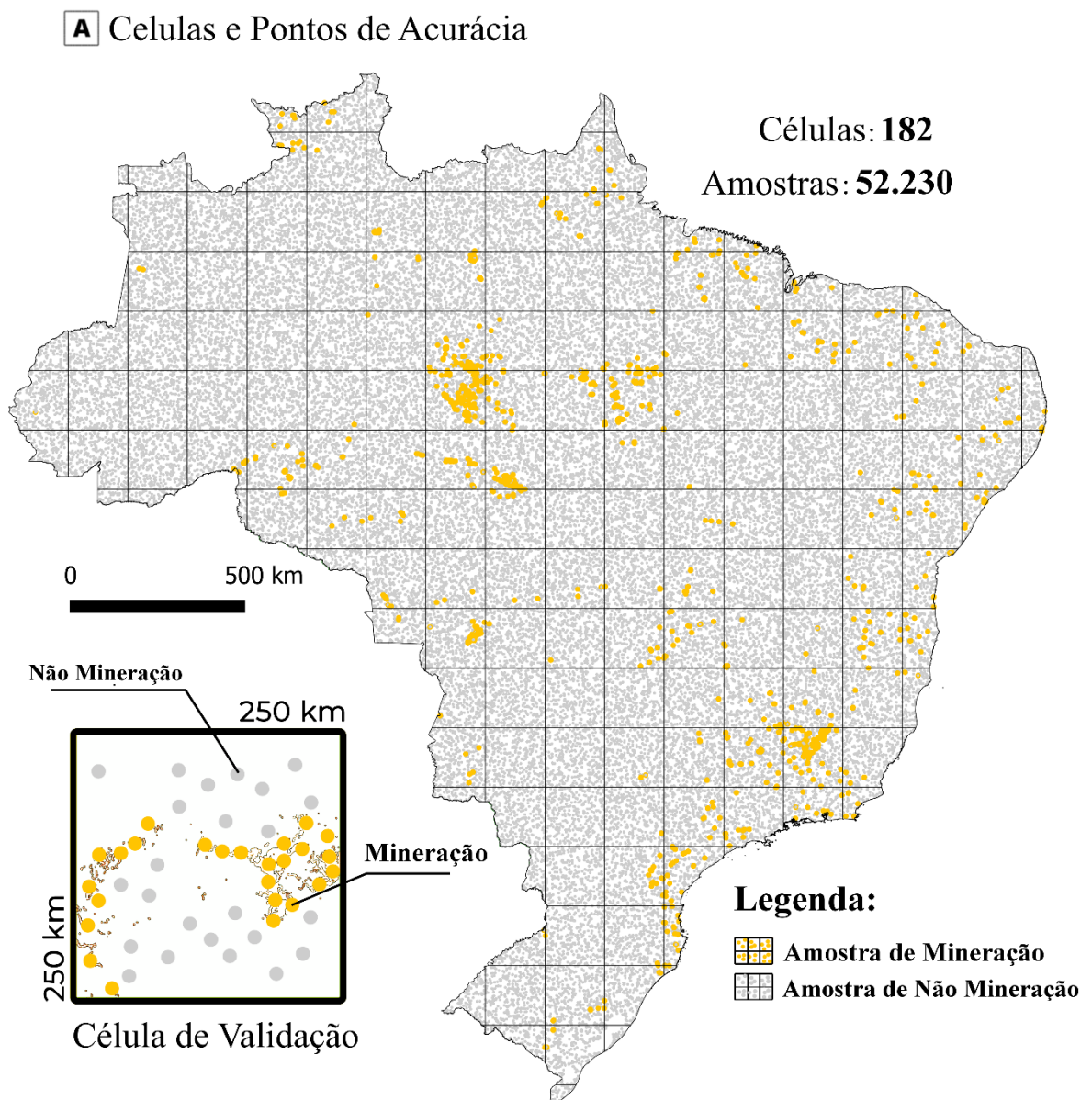


Figura 22 - Estratégia de amostragem aleatória e duplamente estratificada por célula. Fonte: O Autor (2024).

Por seguinte, as amostras são avaliadas individualmente por três especialistas que de posse de mosaicos anuais Landsat e de mosaicos mensais Planet, computam seus votos. Em

seguida, os votos são compilados, e pela maioria de voto a classe é escolhida para cada amostra.

Na Figura 23 são apresentadas séries temporais das métricas de avaliação do dado, as métricas (OLOFSSON *et al.*, 2014) são: acurácia global (O.A), acurácia do produtor (A.P) e acurácia do consumidor (A.C), para cada um dos anos gerados.

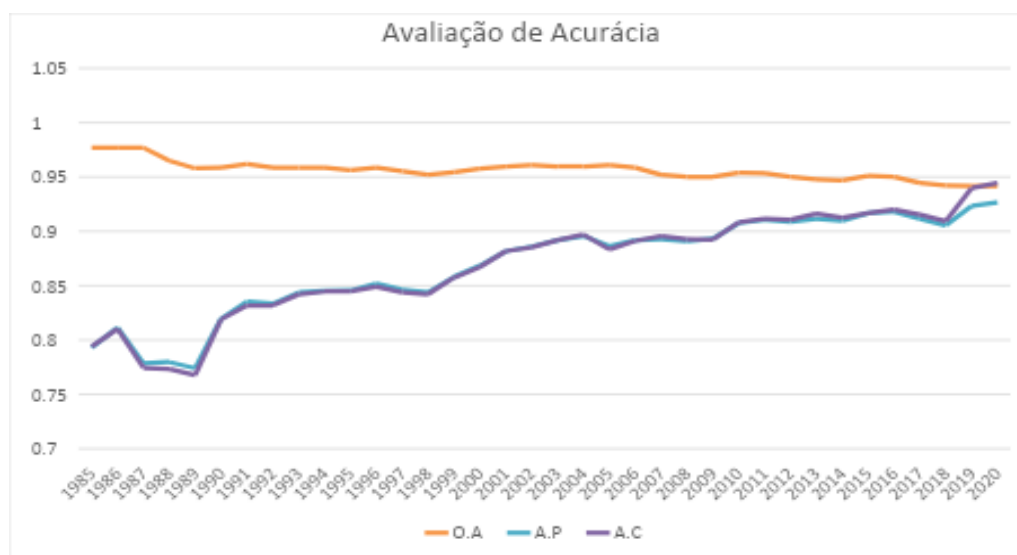
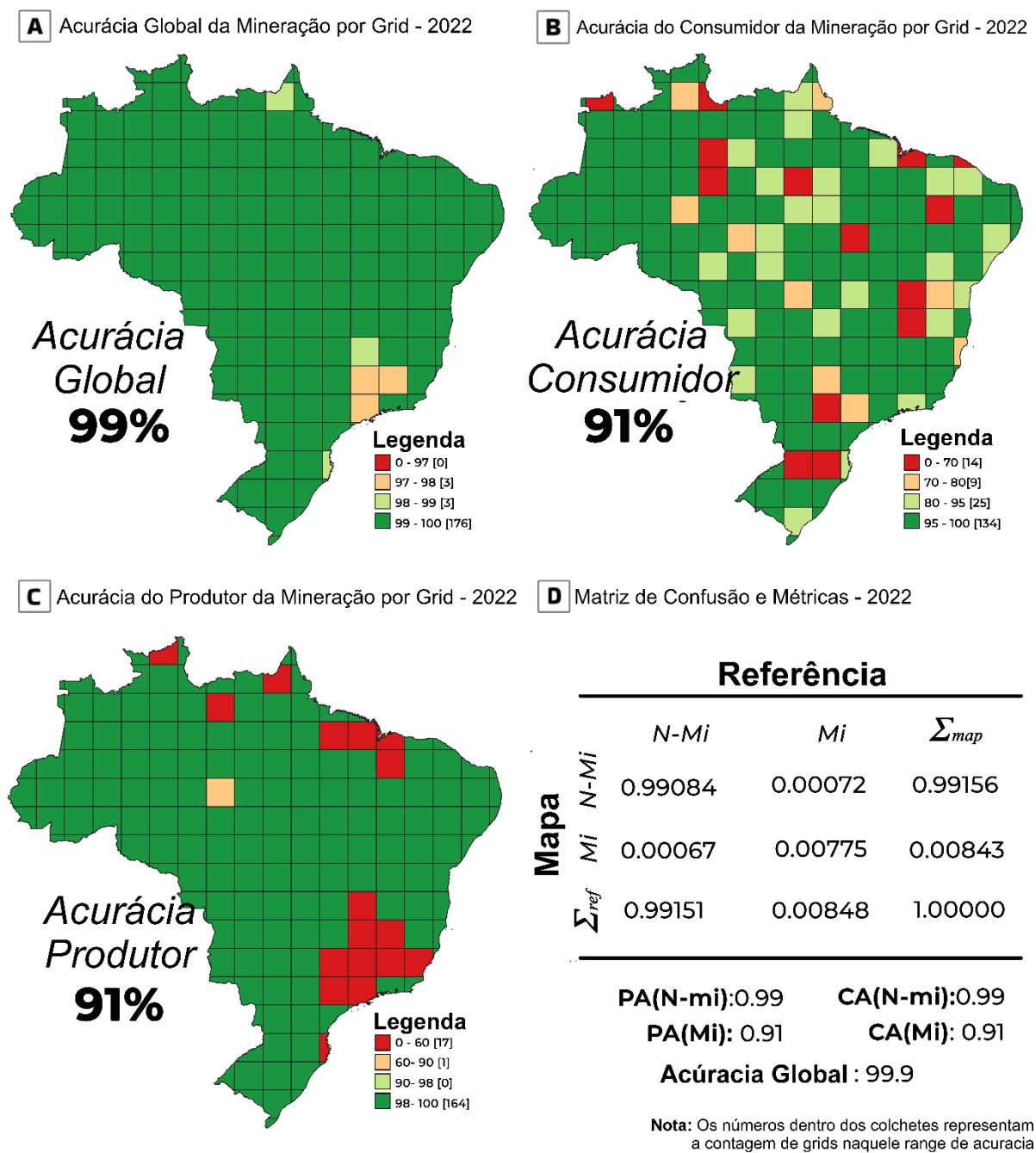


Figura 23- Série Temporal de avaliação da acurácia. Fonte: O Autor (2024).

A Figura 24 apresenta uma análise espacialmente detalhada da performance de um classificador aplicado à detecção de áreas de mineração no Brasil no ano de 2022. A avaliação é segmentada em quatro componentes (A, B, C e D), que, em conjunto, oferecem uma visão abrangente do desempenho do modelo. O mapa A, referente à acurácia geral por célula, revela uma predominância de alta acurácia, com 99% de acurácia geral e 176 grades classificadas na faixa de 99-100%. Esse resultado indica um desempenho robusto do classificador na maior

parte do território nacional, sugerindo uma alta capacidade de discriminação entre áreas de mineração e não-mineração.



classificador. Ambos os mapas apresentam uma acurácia média de 91%, mas com uma variabilidade espacial mais acentuada em comparação ao mapa A. Observa-se uma quantidade significativa de grades com alta acurácia 134 grades entre 95-100% para Consumidor e 164 grades entre 98-100% para Produtor, mas também a presença de regiões com menor performance. As áreas com menor acurácia do consumidor, indicadas em tons de vermelho e laranja, sinalizam locais onde o classificador tende a cometer mais erros de comissão, ou seja, classificar erroneamente áreas como sendo de mineração. De forma análoga, as áreas em vermelho no mapa de acurácia do produtor indicam regiões com maior propensão a erros de omissão, onde áreas de mineração não são corretamente identificadas como tal.

A análise conjunta dos mapas de acurácia do consumidor e do produtor sugere que as regiões Sudeste e Centro-Oeste do Brasil apresentam os maiores desafios para o classificador. Nessas áreas, a menor acurácia pode estar relacionada a diversos fatores, como a maior complexidade ambiental, a presença de características espectrais mais heterogêneas ou a menor disponibilidade de dados de treinamento de alta qualidade. Uma investigação mais aprofundada das características dessas regiões é necessária para compreender as causas específicas da menor performance do classificador e direcionar melhorias futuras no modelo.

O componente D da Figura 24 fornece uma validação quantitativa do desempenho do classificador através da matriz de confusão e das métricas de acurácia do produtor, acurácia do consumidor e acurácia geral. A matriz de confusão confirma a alta acurácia geral (99,9%), detalhando a proporção de classificações corretas e incorretas para as classes de mineração e não-mineração. As métricas de acurácia do produtor e do consumidor corroboram os valores médios observados nos mapas B e C (91% para ambas), reforçando a confiabilidade da análise espacial apresentada. Em suma, a figura oferece uma avaliação abrangente e espacialmente explícita do classificador, permitindo uma compreensão detalhada de seus pontos fortes e fracos, bem como a identificação de áreas que requerem atenção especial para aprimoramento do modelo.

6 Conclusões

6.1 Desempenho e Validação da Metodologia Proposta

O método apresentado neste documento é o primeiro método sistemático para o mapeamento da mineração em todo o Brasil. O método uma vez que tem seu classificador treinado, é livre de interferências humanas, é capaz de processar de forma autônoma mosaicos anuais livres de nuvem e com tempo de processamento diminuído, pela modificação na arquitetura da U-NET.

A acurácia global média do mapa foi de 98%, variando entre 94,16% (2022) e 97,7% (1987). Para avaliar erros específicos, foram utilizadas as métricas de acurácia do produtor e acurácia do consumidor, que mede a probabilidade de um pixel da classe de referência ser corretamente classificado, apresentou uma média de 86,9% de A.P, que mede a probabilidade de um pixel classificado como mineração realmente pertencer a essa classe, teve uma média de 86,97% A.C. A escolha dessas métricas permite avaliar tanto a capacidade do método de identificar corretamente áreas de mineração (A.P.) quanto a confiabilidade da classificação resultante (A.C.).

Esta tese apresenta um método automatizado para mapear a mineração no Brasil a partir de imagens Landsat, utilizando uma rede neural convolucional em formato de U (U-Net). A U-Net modificada e proposta se mostrou 30% menor em quantidade de parâmetros treináveis em comparação com U-Net original. O método proposto foi aplicado a mosaicos anuais Landsat livres de nuvens, abrangendo um período de 37 anos (1985-2022). A acurácia da U-Net foi validada espacialmente por especialistas, obtendo-se uma média de 99% de acurácia geral, 91% de acurácia do produtor e 91% de acurácia do consumidor.

O método proposto é eficaz para o mapeamento da atividade mineradora, fornecendo dados precisos e atualizados para a gestão ambiental e o planejamento territorial. Os resultados demonstram um aumento expressivo da área ocupada pela mineração, que cresceu cerca de 10 vezes entre 1985 e 2022, totalizando 4.500 km². O garimpo foi o subtipo de mineração que mais cresceu percentualmente, passando de 218 km² para 2.627km².

A U-Net modificada e proposta é mais eficiente em termos de recursos computacionais do que a U-Net original, tornando-a mais adequada para o processamento de grandes conjuntos

de dados. A redução de 30% na quantidade de parâmetros treináveis resultou em um tempo de treinamento 60% menor, sem perda significativa de acurácia. Para avaliar as modificações na arquitetura da U-Net, foi realizado um procedimento de testes que consistiu em selecionar aleatoriamente um conjunto de 10 regiões medindo 10km x 10km, construir um dataset de chips para o treinamento das duas U-Nets (modificada e original) e, em seguida, classificar 37 anos de dados dos 10 chips, calculando a área total para a classe de mineração predita. Os resultados mostraram que a U-Net modificada teve um desempenho similar ao da U-Net original, mas com uma redução de 30% na quantidade de parâmetros. A eficiência computacional do modelo proposto é particularmente importante para o processamento de grandes conjuntos de dados, como os utilizados nesta tese, que abrangem todo o território brasileiro ao longo de 37 anos.

O método proposto pode ser usado para monitorar a evolução da área de mineração ao longo do tempo, o que é importante para a identificação de tendências e padrões. Os resultados demonstram um aumento expressivo da área ocupada pela mineração, especialmente pelo garimpo, nas últimas décadas. O monitoramento da evolução da área de mineração ao longo do tempo é crucial para a gestão ambiental e o planejamento territorial, permitindo a identificação de áreas de risco e a adoção de medidas preventivas.

As conclusões de engenharia desta tese contribuem para o desenvolvimento de métodos mais eficientes e precisos para o mapeamento da mineração, o que é importante para a gestão ambiental e o planejamento territorial, além de atingirem um marco inicial para as próximas pesquisas que desejem criar metodologias para o mapeamento sistemático e autônomo de minerações.

6.2 Importância Sócio-política do Dado Gerado

Os 37 anos de mineração, agora disponíveis, são inéditos no mundo científico e como um grande compilado é inédito para a sociedade civil como um todo. Nos mapas, provocações espacialmente apontadas sobre a evolução das minerações de garimpo são inéditas, e pela primeira vez se pode inferir sobre padrões de expansão e como as regiões de seu entorno são impactadas.

Com apoio dos dados produzidos aqui nesta tese, é possível calcular as transições desde 1985 a 2022, habilitando a percepção de que a mineração transformou 100.000ha de vegetação nativa em áreas de explorações, o equivalente a ~82.500 campos de futebol padrão FIFA, destas modificações 87% estão no estado do Pará o qual possui em seu Sudeste a maior concentração de garimpos de toda Amazônia Legal, cerca de 70.000ha de explorações provavelmente ilegais estão no perímetro dos rios Tapajós, Crepori e seus afluentes.

Agora, com os 37 mapas de mineração gerados, a possibilidade de uso é grande, seja para informação da sociedade civil, para o cálculo do dano ambiental antecedente, ou para o entendimento dos padrões de comportamento dos *hotspots* ainda ativos.

Em termos acadêmicos, classificamos o garimpo como ASM ou mineração artesanal de ouro em pequena escala. No entanto, somente em 2022, a área de garimpo aumentou 341 km² e ultrapassou a área de mineração industrial em 848 km² (uma área de mineração industrial de 1779 km² e uma área de garimpo de 2627 km²). No mesmo ano, dentro da Terra Indígena Kayapó, onde a atividade de mineração de qualquer tipo é restrita, um único conglomerado de garimpo cobria uma área de 50 km², alcançando extensões lineares de 20 km apresentado na Figura 25. Assim, o conceito de pequena escala não se aplica mais, uma vez que tais extensões e áreas não podem representar medidas de pequena escala. Da mesma forma, não há nada de artesanal associado ao tamanho da força de trabalho e ao volume de transporte de várias toneladas de sedimentos escavados por meio do uso de dezenas, às vezes centenas, de retroescavadeiras e escavadeiras hidráulicas, que muitas vezes são combinadas com o uso de helicópteros, pistas de pouso clandestinas e balsas de dragagem para aumentar a produção e a capacidade de vazão dos garimpos amazônicos. Tais dimensões e velocidades de expansão são

incompatíveis com o suposto conceito artesanal e de pequena escala.

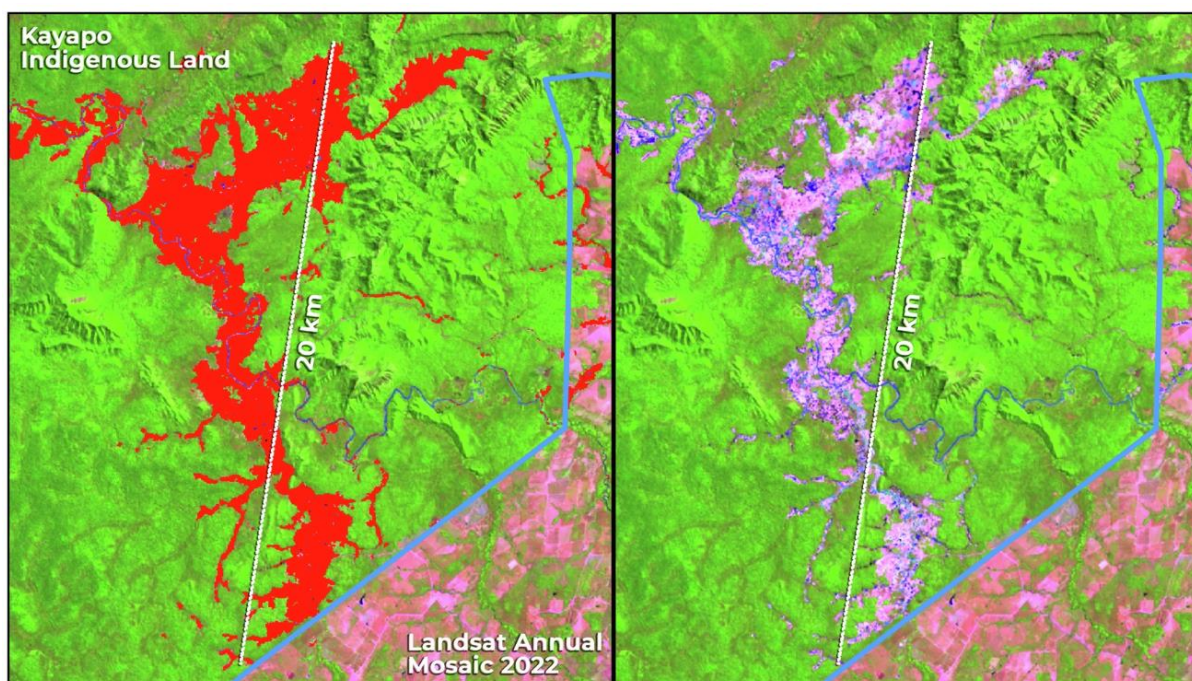


Figura 25 - Terra Indígena Kayapó e garimpo em Vermelho, maior garimpo de extensão em linha reta, são 20km. Fonte: O Autor (2024).

Em todas as seções territoriais analisadas, seja dentro de áreas restritas ou fora, a área de garimpo aumentou substancialmente ao longo dos anos. No entanto, a progressão do aumento nos últimos anos é preocupante. De 2018 a 2022, a área de garimpo se expandiu em quase 1000 km², representando uma expansão de curto prazo de 60% em apenas cinco anos Figura 20. A expansão recente também se reflete dentro de TIs e UCs. Na última década, para TIs combinadas com UCs, a área de garimpo foi 440% maior em 2022 do que em 2018 Figura 21. Dentro apenas de TIs, a área em 2022 foi 250% maior do que em 2018, variando de 100 a 250 km² Figura 21. Dentro de UCs, a extensão da extração foi 190% maior em 2022 do que em 2018, expandindo de 410 para 780 km² Figura 21. No geral, o aumento dentro de TIs e UCs atingiu 520 km² em apenas cinco anos. Essa expansão rápida e espacialmente persistente, dentro e fora de áreas restritas à mineração, sejam TIs ou UCs, indica a falta de controle do Brasil sobre a mineração ilegal na Amazônia. Notavelmente, TIs e UCs são as unidades territoriais que enfrentam uma necessidade urgente de assistência 59,60. Embora existam 48 UCs com sinais de extração de garimpo, apenas cinco UCs representam 85,5% da área total de garimpo dentro dessas unidades: APA do Tapajós, FLONA do Amana, ESEC Juami-Japurá, FLONA do Crepori e PARNA do Rio Novo. Da mesma forma, foram detectadas 22 TIs com

sinais de atividade de garimpo. No entanto, três TIs sozinhas - Kayapó, Munduruku e Yanomami - representaram mais de 90% da área total de mineração nessas regiões protegidas Figura 21.

Desde 2000, o mercado internacional de ouro tem mostrado sinais de crescimento persistente. Concomitantemente, mudanças na legislação brasileira facilitaram a expansão do garimpo 61. Em 2008, a Lei nº 11.685/2008 62 criou o Estatuto do Garimpeiro, diferenciando o garimpo da mineração industrial, considerando a tecnologia de pequena escala do primeiro e a possibilidade de ser praticado individualmente ou coletivamente por meio de cooperativas. Essa lei, baseada mais uma vez no argumento da pequena escala, desregulamentou os garimpeiros da necessidade de apresentar uma autorização de pesquisa mineral (uma autorização da ANM anterior à abertura da mina, que visa demonstrar a viabilidade econômica e ambiental da operação proposta). Hoje, esse mecanismo regulatório se aplica apenas a locais de mineração industrial, ao contrário das portarias formalmente expressas na Resolução CONAMA nº 237/1997 63. Além disso, em 2013, a Lei nº 12.844/2013 64 expandiu a desregulamentação da mineração ao presumir que as pessoas jurídicas que compram e vendem ouro agem de boa fé. Do ponto de vista jurídico, a lei da boa-fé permitiu a presunção de legalidade dos documentos de mineração, o que, por sua vez, excluiu a inspeção rotineira de sua legitimidade e veracidade. Dez anos depois, em 5 de maio de 2023, o Supremo Tribunal Federal derrubou a validade dessa lei.

Em 2019, 2020 e 2022, surgiram três projetos de lei, que ainda aguardam uma decisão definitiva, mas demonstram que a desregulamentação da atividade de garimpo ainda está ativa em uma parte do corpo político brasileiro que busca expandir ainda mais a permissividade da mineração na Amazônia. Os Projetos de Lei nº 5822/2019 65, 191/2020 66 e 571/2022 67 podem possibilitar a implementação de empreendimentos de mineração dentro de florestas nacionais (FLONA), estabelecer condições para a pesquisa e exploração de recursos minerais e hidrocarbonetos em TIs e permitir que o Presidente da República libere a extração mineral em qualquer área do país em tempos de crise econômica, incluindo TIs e UCs.

Um número enorme de amostras de validação apoiou os achados estatísticos aqui relatados. Foram mais de 52 mil amostras validadas por humanos, distribuídas ao longo de nove anos diferentes da série temporal (de 1985 a 2022), nas quais foram alcançadas métricas de erro razoavelmente favoráveis, considerando a característica rara da classe de interesse (OA > 98%, PA > 65% e CA > 90%). Por sua vez, os erros de omissão foram mais abundantes do

que os erros de comissão, o que é um aspecto positivo do ponto de vista do usuário do mapa, pois revela que cada menção a um local de mineração no mapa geralmente corresponde à realidade da superfície da Terra. No entanto, como indicador socioambiental, alguma mineração ainda é indetectada, exacerbando a gravidade desse cenário desastroso. No entanto, as altas taxas de acurácia obtidas nesta extensa análise de erros confirmam a eficácia geral e a adequação do método baseado em IA adotado, ao mesmo tempo em que indicam que melhorias locais de treinamento ainda são necessárias, particularmente considerando os três primeiros anos da série temporal, a saber, 1985, 1990 e 1995. Por sua vez, as métricas de erro não podem servir como garantias de qualidade.

As métricas de erro/acurácia são mecanismos para encontrar, entender e corrigir uma não conformidade categórica, seja no tempo ou no espaço, apoiando assim a produção futura de melhores versões de representações da superfície da Terra. Devido à natureza rara do alvo de interesse (considerando a escala continental do Brasil), sem a espacialização local e as abordagens ponderadas aqui aplicadas, as métricas de erro seriam rapidamente infladas ou deflacionadas e, portanto, perderiam o conceito central de uso como uma ferramenta analítica que permite um melhor ajuste entre a representação cartográfica e a realidade. A análise de erros nos permitiu reduzir localmente as inconsistências espaciais e temporais, que estarão disponíveis publicamente na Coleção 9 da iniciativa MapBiomias (MAPBIOMAS, 2021).

As características técnicas da metodologia adotada limitam sua capacidade de detecção e realizações analíticas. Nesse sentido, a atual arquitetura estrutural brasileira e as técnicas relacionadas ao sensoriamento remoto restringem nossa pesquisa. Estruturalmente, como é impraticável obter licenças ambientais, foi reconhecido que toda licença de mineração presente na plataforma SIGMINE também inclui licenciamento ambiental prévio. Essa condição pode precisar ser verificada posteriormente. No entanto, essa tarefa é intransponível, pois o país ainda precisa estabelecer uma base de dados pública e centralizada de licenças ambientais. Do ponto de vista espacial, como em qualquer análise baseada em sensoriamento remoto, a realidade da superfície da Terra pode diferir de sua categorização cartográfica. Nossos resultados foram baseados em dados Landsat (30 m), portanto, locais não detectados são, sem dúvida, menores do que a resolução de detecção do Landsat, apontando para áreas de mineração maiores do que as relatadas aqui. Por outro lado, alguns erros de comissão ainda existem, conforme relatado na seção Avaliação de Erros. Por fim, existem ilegalidades latentes nas atividades de garimpo que não são espaciais nem documentadas e, portanto, estão fora do

escopo do nosso método, como a quantificação do uso irregular de mercúrio e arsênico na amalgamação de ouro e outros metais, a verificação das condições de trabalho associadas, as relações econômicas e logísticas entre o garimpo e as atividades de tráfico de drogas e o mecanismo pelo qual grandes quantidades de ouro ilegal são vendidas e compradas nos mercados formais nacionais e internacionais (lavagem de ouro).

6.3 Trabalhos Futuros

O trabalho realizado nessa tese, compreendeu o tempo de um único doutorado, perante a ciência e o desenvolvimento tecnológica esse tempo foi mínimo e, portanto, diversos pontos podem ser melhorados através de contribuições científicas futuras.

Na área técnica, a integração de outras fontes de dados, de satélites e inputs outros que nos permitam enxergar o mundo com mais detalhamento visual, a exemplo dos satélites Dove da Planet e os satélites Sentinel da ESA. Ainda tecnicamente a metodologia aqui construída se prova eficiente e inovadora para apontar minerações ocorridas no período de um ano, contudo em nada versa sobre o problema da atividade em tempo real, mecanismos outros podem ser explorados para a construção de mecanismos de monitoramento do território em busca de minerações a céu aberto.

A mineração como atividade humana tem seu modo de operação sendo modificado ao longo da história, a mineração de garimpo que outrora era considerada de pequena escala hoje possui cavas com mais de 20km em linha reta, e modalidades novas como as balsas garimpeiras ocorrem em regiões de mineração de aluvião como o rio madeira, e não existem ainda sistemas capazes de detectar a atividade nessa modalidade.

A busca por segmentadores semânticos melhores ou por melhores técnicas de simplificação da rede como o uso de mecanismos de *Knowledge Distillation* (HINTON, G.; VINYALS; DEAN, 2015) e *Model Compression* (DANTAS *et al.*, 2024) podem reforçar ainda mais a característica de democratização do modelo presente nessa tese e diminuir o esforço computacional necessário para a detecção da atividade garimpeira.

A metodologia que foi desenvolvida aqui também pode ser utilizada em outros alvos do sensoriamento remoto, testes em novos alvos podem nos permitir revelar novas classes a sociedade civil e contribuir para o mapeamento de novas atividades.

7 Referências

ADAMI, M. *et al.* Remote sensing time series to evaluate direct land use change of recent expanded sugarcane crop in Brazil. *Sustainability*, v. 4, n. 4, p. 574–585, 2012.

ADEGUN, A. A.; VIRIRI, S.; TAPAMO, J. R. Review of deep learning methods for remote sensing satellite images classification: experimental survey and comparative analysis. *Journal of Big Data*, v. 10, n. 1, 1 dez. 2023a. Acesso em: 22 abr. 2025.

ADEGUN, A. A.; VIRIRI, S.; TAPAMO, J. R. Review of deep learning methods for remote sensing satellite images classification: experimental survey and comparative analysis. *Journal of Big Data*, v. 10, n. 1, p. 1–24, 1 dez. 2023b. Disponível em: <<https://journalofbigdata.springeropen.com/articles/10.1186/s40537-023-00772-x>>. Acesso em: 22 jan. 2025.

ALTAREZ, R. D. Faster R-CNN, RetinaNet and Single Shot Detector in different ResNet backbones for marine vessel detection using cross polarization C-band SAR imagery. *Remote Sensing Applications: Society and Environment*, v. 36, p. 101297, 1 nov. 2024. Acesso em: 22 abr. 2025.

ARTHUR, D.; VASSILVITSKII, S. k-means++: The Advantages of Careful Seeding. 2007. Acesso em: 20 jan. 2025.

ASNER, G. P. *et al.* Elevated rates of gold mining in the Amazon revealed through high-resolution monitoring. *Proceedings of the National Academy of Sciences*, v. 110, n. 46, p. 18454–18459, 2013. Disponível em: <<http://www.pnas.org/cgi/doi/10.1073/pnas.1318271110>>.

BAATZ, M.; SCHÄPE, A. Multiresolution Segmentation : an optimization approach for high quality multi-scale image segmentation. *Angewandte Geographische Informations-Verarbeitung*, v. XII, n. 1, p. 12–23, 2000. Acesso em: 15 fev. 2025.

BAGCI, R. Ş.; ACAR, E.; TÜRK, Ö. Identification of cotton and corn plant areas by employing deep transformer encoder approach and different time series satellite images: A case study in Diyarbakir, Turkey. *Computers and Electronics in Agriculture*, v. 209, p. 107838, 1 jun. 2023. Acesso em: 22 abr. 2025.

BALAGURUNATHAN, Y. *et al.* Semantic Guidance Fusion Network for Cross-Modal Semantic Segmentation. *Sensors 2024, Vol. 24, Page 2473*, v. 24, n. 8, p. 2473, 12 abr. 2024. Disponível em: <<https://www.mdpi.com/1424-8220/24/8/2473/htm>>. Acesso em: 14 jan. 2025.

BENGIO, Y.; SIMARD, P.; FRASCONI, P. Learning Long-Term Dependencies with Gradient Descent is Difficult. *IEEE Transactions on Neural Networks*, v. 5, n. 2, p. 157–166, 1994. Acesso em: 20 jan. 2025.

BERZAS NEVADO, J. J. *et al.* Mercury in the Tapajós River basin, Brazilian Amazon: A review. *Environment International*. [S.l: s.n.], 2010

BESERRA, A. A. R.; GOULARTE, R. Multimodal early fusion operators for temporal video scene segmentation tasks. *Multimedia Tools and Applications*, v. 82, n. 20, p. 31539–31556, 1 ago. 2023. Disponível em: <<https://repositorio.usp.br/item/003126935>>. Acesso em: 14 jan. 2025.

BOCK, H.-H. Clustering Methods: A History of k-Means Algorithms. v. 1, p. 161–172, 2007. Disponível em: <https://link.springer.com/chapter/10.1007/978-3-540-73560-1_15>. Acesso em: 15 fev. 2025.

BOSTON, T. *et al.* Comparing CNNs and Random Forests for Landsat Image Segmentation Trained on a Large Proxy Land Cover Dataset. *Remote Sensing*, v. 14, n. 14, p. 3396, 1 jul. 2022. Disponível em: <<https://www.mdpi.com/2072-4292/14/14/3396/htm>>. Acesso em: 20 jan. 2025.

BRAR, K. K. *et al.* Image segmentation review: Theoretical background and recent advances. *Information Fusion*, v. 114, p. 102608, 1 fev. 2025. Acesso em: 21 jan. 2025.

BREIMAN, L. Random forest. *Machine Learning*, v. 45, n. 5, p. 1–35, 1999.

CAI, Z. *et al.* A cost-effective and robust mapping method for diverse crop types using weakly supervised semantic segmentation with sparse point samples. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 218, p. 260–276, 1 dez. 2024. Acesso em: 22 abr. 2025.

CANTORNA, D. *et al.* Oil spill segmentation in SAR images using convolutional neural networks. A comparative analysis with clustering and logistic regression algorithms. *Applied Soft Computing*, v. 84, p. 105716, 1 nov. 2019. Acesso em: 14 jan. 2025.

CHANDER, G.; MARKHAM, B. L.; HELDER, D. L. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, v. 113, n. 5, 2009.

CORTINHAS FERREIRA NETO, L. *et al.* Uncontrolled Illegal Mining and Garimpo in the Brazilian Amazon. *Nature Communications* 2024 15:1, v. 15, n. 1, p. 1–15, 13 nov. 2024. Disponível em: <<https://www.nature.com/articles/s41467-024-54220-2>>. Acesso em: 13 jan. 2025.

CRESPO-LOPEZ, M. E. *et al.* Mercury: What can we learn from the Amazon? *Environment International*. [S.l.: s.n.], 2021

DANTAS, P. V. *et al.* A comprehensive review of model compression techniques in machine learning. *Applied Intelligence* 2024 54:22, v. 54, n. 22, p. 11804–11844, 2 set. 2024. Disponível em: <<https://link.springer.com/article/10.1007/s10489-024-05747-w>>. Acesso em: 16 jan. 2025.

DE MORAES NOVO, E. M. L. *Sensoriamento remoto: principios e aplicações*. [S.l.]: Blucher, 2010. Disponível em: <<https://books.google.com.br/books?id=M6kOuQAACAAJ>>.

DERRODE, S.; PIECZYNSKI, W. Unsupervised classification using hidden Markov chain with unknown noise copulas and margins. *Signal Processing*, v. 128, p. 8–17, 1 nov. 2016. Acesso em: 14 jan. 2025.

DING, L.; TANG, H.; BRUZZONE, L. LANet: Local Attention Embedding to Improve the Semantic Segmentation of Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, v. 59, n. 1, p. 426–435, 1 jan. 2021. Acesso em: 22 jan. 2025.

DINIZ, C. *et al.* A large-scale deep-learning approach for multi-temporal aqua and salt-culture mapping. *Remote Sensing*, v. 13, n. 8, 2021.

DINIZ, C. G. *et al.* DETER-B: The New Amazon Near Real-Time Deforestation Detection System. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 8, n. 7, p. 3619–3628, 2015.

DINIZ, C. G. *Três décadas de mudanças na planície costeira brasileira: O status dos manguezais, da aquicultura e salicultura a partir de séries temporais Landsat e técnicas de aprendizado de máquina*. 2020. 1–120 f. Thesis – Universidade Federal do Pará, Belém, 2020. Disponível em: <<https://repositorio.ufpa.br/jspui/handle/2011/13504>>. Acesso em: 20 jan. 2025.

DIXIT, M.; CHAURASIA, K.; KUMAR MISHRA, V. Dilated-ResUnet: A novel deep learning architecture for building extraction from medium resolution multi-spectral satellite imagery. *Expert Systems with Applications*, v. 184, p. 115530, 1 dez. 2021. Acesso em: 22 abr. 2025.

EBRAHIMI, S.; KUMAR, S. Semantic segmentation for simultaneous crop and land cover land use classification using multi-temporal Landsat imagery. *Remote Sensing Applications: Society and Environment*, v. 37, p. 101505, 1 jan. 2025. Acesso em: 22 abr. 2025.

ELMAN, J. L. Finding structure in time. *Cognitive Science*, v. 14, n. 2, p. 179–211, 1 abr. 1990. Acesso em: 20 jan. 2025.

EMEK SOYLU, B. *et al.* Deep-Learning-Based Approaches for Semantic Segmentation of Natural Scene Images: A Review. *Electronics 2023, Vol. 12, Page 2730*, v. 12, n. 12, p. 2730, 19 jun. 2023. Disponível em: <<https://www.mdpi.com/2079-9292/12/12/2730/htm>>. Acesso em: 14 jan. 2025.

EMERY, W.; CAMPS, A. The History of Satellite Remote Sensing. In: WILLIAM EMERY, A. C. (Org.). *Introduction to Satellite Remote Sensing*. 1. ed. [S.l.]: Elsevier, 2017.

v. 1. p. 1–42. Disponível em:
<<https://linkinghub.elsevier.com/retrieve/pii/B9780128092545000014>>.

FERRARI, F. *et al.* Fusing Sentinel-1 and Sentinel-2 Images for Deforestation Detection in the Brazilian Amazon Under Diverse Cloud Conditions. *IEEE Geoscience and Remote Sensing Letters*, v. 20, 2023. Acesso em: 22 abr. 2025.

FERRARI, L. *et al.* Integrating EfficientNet into an HAFNet Structure for Building Mapping in High-Resolution Optical Earth Observation Data. *Remote Sensing 2021, Vol. 13, Page 4361*, v. 13, n. 21, p. 4361, 29 out. 2021. Disponível em: <<https://www.mdpi.com/2072-4292/13/21/4361/htm>>. Acesso em: 22 jan. 2025.

FORGY, E. W. Cluster analysis of multivariate data : efficiency versus interpretability of classifications. *Biometrics*, 1965. Acesso em: 14 jan. 2025.

FU, B. *et al.* Comparison of optimized object-based RF-DT algorithm and SegNet algorithm for classifying Karst wetland vegetation communities using ultra-high spatial resolution UAV data. *International Journal of Applied Earth Observation and Geoinformation*, v. 104, p. 102553, 15 dez. 2021. Acesso em: 22 abr. 2025.

FU, Y. *et al.* Mapping impervious surfaces in town-rural transition belts using China's GF-2 imagery and object-based deep CNNs. *Remote Sensing*, v. 11, n. 3, 2019.

GAO, B. C. NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, v. 58, n. 3, p. 257–266, 1996.

GAO, R. Rethinking Dilated Convolution for Real-time Semantic Segmentation. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, v. 2023- June, p. 4675–4684, 2023. Acesso em: 14 jan. 2025.

GAUTAM, S.; MATHURIA, T.; MEENA, S. Image Segmentation for Self-Driving Car. *2022 2nd International Conference on Intelligent Technologies, CONIT 2022*, 2022. Acesso em: 14 jan. 2025.

GER, F. A.; SCHMIDHUBER, J.; CUMMINS, F. Learning to forget: Continual prediction with LSTM. *IEE Conference Publication*, v. 2, n. 470, p. 850–855, 1999. Acesso em: 20 jan. 2025.

GYAWALI, A. *et al.* From simple linear regression to machine learning methods: Canopy cover modelling of a young forest using planet data. *Ecological Informatics*, v. 82, p. 102706, 1 set. 2024. Acesso em: 22 abr. 2025.

HAN, H. *et al.* Optimizing image captioning: The effectiveness of vision transformers and VGG networks for remote sensing. *Big Data Research*, v. 37, p. 100477, 28 ago. 2024. Acesso em: 22 abr. 2025.

HAO, S.; ZHOU, Y.; GUO, Y. A Brief Survey on Semantic Segmentation with Deep Learning. *Neurocomputing*, v. 406, p. 302–321, 17 set. 2020. Acesso em: 14 jan. 2025.

HARTIGAN, J. A.; WONG, M. A. Algorithm AS 136: A K-Means Clustering Algorithm. *Applied Statistics*, v. 28, n. 1, p. 100, 1979. Acesso em: 20 jan. 2025.

HASLAM, P. A.; ARY TANIMOUNE, N. The Determinants of Social Conflict in the Latin American Mining Sector: New Evidence with Quantitative Data. *World Development*, v. 78, 2016.

HATAMIZADEH, A. *et al.* UNetFormer: A Unified Vision Transformer Model and Pre-Training Framework for 3D Medical Image Segmentation. 1 abr. 2022. Disponível em: <<https://arxiv.org/abs/2204.00631v2>>. Acesso em: 15 jan. 2025.

HELWEGE, A. Challenges with resolving mining conflicts in Latin America. *Extractive Industries and Society*, v. 2, n. 1, 2015.

HINTON, G. E.; OSINDERO, S.; TEH, Y.-W. A Fast Learning Algorithm for Deep Belief Nets. *Neural Computation*, v. 18, n. 7, p. 1527–1554, jul. 2006. Disponível em: <<https://direct.mit.edu/neco/article/18/7/1527-1554/7065>>. Acesso em: 14 jan. 2025.

HINTON, G.; VINYALS, O.; DEAN, J. Distilling the Knowledge in a Neural Network. 9 mar. 2015. Disponível em: <<https://arxiv.org/abs/1503.02531v1>>. Acesso em: 16 jan. 2025.

HUANG, Y.; CHEN, Y. Autonomous Driving with Deep Learning: A Survey of State-of-Art Technologies. *Computer Vision and Pattern Recognition*, 2020.

HUVAL, B. *et al.* An Empirical Evaluation of Deep Learning on Highway Driving. 7 abr. 2015. Disponível em: <<https://arxiv.org/abs/1504.01716v3>>. Acesso em: 20 jan. 2025.

HYNDMAN, R. J.; KOEHLER, A. B. Another look at measures of forecast accuracy. *International Journal of Forecasting*, v. 22, n. 4, p. 679–688, 1 out. 2006. Acesso em: 22 jan. 2025.

IKOTUN, A. M. *et al.* K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, v. 622, p. 178–210, 1 abr. 2023. Acesso em: 23 jan. 2025.

IQBAL, S. *et al.* On the Analyses of Medical Images Using Traditional Machine Learning Techniques and Convolutional Neural Networks. *Archives of Computational Methods in Engineering* 2023 30:5, v. 30, n. 5, p. 3173–3233, 4 abr. 2023. Disponível em: <<https://link.springer.com/article/10.1007/s11831-023-09899-9>>. Acesso em: 14 jan. 2025.

ISMAIL FAWAZ, H. *et al.* Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery*, v. 33, n. 4, 2019.

IUBEL, A. F. Terras de Ouro: Narrativas e experiências indígenas e não indígenas acerca do garimpo de ouro na Amazônia Brasileira. *Anuário Antropológico*, n. v.45 n.1, 2020.

JAIN, A. K. Data clustering: 50 years beyond K-means. *Pattern Recognition Letters*, v. 31, n. 8, p. 651–666, 1 jun. 2010. Acesso em: 20 jan. 2025.

JANNAT, F. E.; WILLIS, A. R. Improving Classification of Remotely Sensed Images with the Swin Transformer. *Conference Proceedings - IEEE SOUTHEASTCON*, v. 2022-March, p. 611–618, 2022. Acesso em: 22 jan. 2025.

JENSEN, J. R. Remote Sensing of the Environment: An Earth Resource Perspective, 2nd Ed. *Prentice Hall, Upper Saddle River, NJ*, 2009.

JENSEN, JOHN R. *Remote sensing of the environment: an earth resource perspective second edition*. 3. ed. New Jersey: Pearson Prentice Hall, 2014. v. 1.

KARIM, M. *et al.* ViT-ChangeFormer: A deep learning approach for cropland abandonment detection in lahore, Pakistan using Landsat-8 and Sentinel-2 data. *Remote Sensing Applications: Society and Environment*, v. 37, p. 101468, 1 jan. 2025. Acesso em: 22 abr. 2025.

KRIZHEVSKY, A.; SUTSKEVER, I.; GEOFFREY E., H. ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems 25 (NIPS2012)*, p. 1–9, 2012. Disponível em: <<https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>>.

KRZANOWSKI, W. J.; LAI, Y. T. A Criterion for Determining the Number of Groups in a Data Set Using Sum-of-Squares Clustering. *Biometrics*, v. 44, n. 1, p. 23, mar. 1988. Acesso em: 15 jan. 2025.

LAO, C. *et al.* Monitoring soil salinity in coastal wetlands with Sentinel-2 MSI data: Combining fractional-order derivatives and stacked machine learning models. *Agricultural Water Management*, v. 306, p. 109147, 20 dez. 2024. Acesso em: 22 abr. 2025.

LARRETA, E. R. “Gold is illusion.” *The Garimpeiros of Tapajos Valley in the Brazilian Amazonia*. 1. ed. Stockholm: Stockholm Universite, 2002. . Acesso em: 7 dez. 2021.

LARY, D. J. *et al.* Machine learning in geosciences and remote sensing. *Geoscience Frontiers*, v. 7, n. 1, 2016.

LEÃO, R.; RABELO, R. SUMÁRIO EXECUTIVO A EXTENSÃO DA CADEIA PRODUTIVA DA ECONOMIA MINERAL NO PIB BRASILEIRO. 2023. Disponível em: <<http://dx.doi.org/10.38116/td2950-port>>. Acesso em: 23 jan. 2025.

LECUN, Y. *et al.* Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation*, v. 1, n. 4, p. 541–551, dez. 1989. Acesso em: 20 jan. 2025.

LECUN, Y. *et al.* Gradient-Based Learning Applied to Document Recognition. 1998. Acesso em: 20 jan. 2025.

LEI, S.; SHI, Z.; ZOU, Z. Coupled Adversarial Training for Remote Sensing Image Super-Resolution. *IEEE Transactions on Geoscience and Remote Sensing*, v. 58, n. 5, 2020.

LI, C. *et al.* Triple generative adversarial nets. 2017, [S.l: s.n.], 2017.

LI, H. *et al.* Global and Local Contrastive Self-Supervised Learning for Semantic Segmentation of HR Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, v. 60, 19 jun. 2021. Disponível em: <<http://arxiv.org/abs/2106.10605>>. Acesso em: 22 jan. 2025.

LI, J. *et al.* A review of remote sensing image segmentation by deep learning methods. *International Journal of Digital Earth*, v. 17, n. 1, 31 dez. 2024. Disponível em: <<https://www.tandfonline.com/doi/abs/10.1080/17538947.2024.2328827>>. Acesso em: 22 jan. 2025.

LI, Q. *et al.* Three-dimensional dynamic monitoring of crevasses based on deep learning and surface elevation reconstruction methods. *International Journal of Applied Earth Observation and Geoinformation*, v. 132, p. 104017, 1 ago. 2024. Acesso em: 22 abr. 2025.

LI, WEI *et al.* An attention-enhanced spatial–temporal high-resolution network for irrigated area mapping using multitemporal Sentinel-2 images. *International Journal of*

Applied Earth Observation and Geoinformation, v. 132, p. 104040, 1 ago. 2024. Acesso em: 22 abr. 2025.

LI, WENYUAN; CHEN, H.; SHI, Z. Semantic Segmentation of Remote Sensing Images with Self-Supervised Multitask Representation Learning. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, v. 14, p. 6438–6450, 2021. Acesso em: 22 jan. 2025.

LI, X. *et al.* Transformer-Based Visual Segmentation: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19 abr. 2023. Disponível em: <<https://arxiv.org/abs/2304.09854v4>>. Acesso em: 14 jan. 2025.

LI, Y. *et al.* Customer segmentation using K-means clustering and the adaptive particle swarm optimization algorithm. *Applied Soft Computing*, v. 113, p. 107924, 1 dez. 2021. Acesso em: 23 jan. 2025.

LITJENS, G. *et al.* A survey on deep learning in medical image analysis. *Medical Image Analysis*. [S.l: s.n.], , 2017

LOBO, F. DE L. *et al.* Mapping mining areas in the Brazilian amazon using MSI/Sentinel-2 imagery (2017). *Remote Sensing*, 2018.

LOBO, F. L.; COSTA, M. P. F.; NOVO, E. M. L. M. Time-series analysis of Landsat-MSS/TM/OLI images over Amazonian waters impacted by gold mining activities. *Remote Sensing of Environment*, v. 157, p. 170–184, 2015. Disponível em: <<http://dx.doi.org/10.1016/j.rse.2014.04.030>>.

LU, H. *et al.* Landslides Information Extraction Using Object-Oriented Image Analysis Paradigm Based on Deep Learning and Transfer Learning. *Remote Sensing* 2020, Vol. 12, Page 752, v. 12, n. 5, p. 752, 25 fev. 2020. Disponível em: <<https://www.mdpi.com/2072-4292/12/5/752/htm>>. Acesso em: 14 jan. 2025.

LV, Z. *et al.* Parallel K-Means Clustering of Remote Sensing Images Based on MapReduce. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 6318 LNCS, n. M4D, p. 162–170, 2010. Disponível em: <https://link.springer.com/chapter/10.1007/978-3-642-16515-3_21>. Acesso em: 23 jan. 2025.

MA, L. *et al.* Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*. [S.l: s.n.], , 2019

MA, X. *et al.* MANet: Fine-Tuning Segment Anything Model for Multimodal Remote Sensing Semantic Segmentation. 15 out. 2024. Disponível em: <<https://arxiv.org/abs/2410.11160v1>>. Acesso em: 22 jan. 2025.

MACHADO, I. F.; SILVA, S. F. 500 years of mining in Brazil: A brief review. *Resources Policy*, v. 27, n. 1, p. 9–24, 2001. Disponível em: <https://www.researchgate.net/publication/222550142_500_years_of_mining_in_Brazil_A_brief_review>. Acesso em: 13 jan. 2025.

MACQUEEN, J. Some methods for classification and analysis of multivariate observations. 1967. Acesso em: 20 jan. 2025.

MADASA, A.; ORIMOLOYE, I. R.; OLOLADE, O. O. Application of geospatial indices for mapping land cover/use change detection in a mining area. *Journal of African Earth Sciences*, v. 175, 2021.

MAPBIOMAS. *Mapbiomas* 2021. Disponível em: <<https://plataforma.brasil.mapbiomas.org/>>.

Markov Chains and Queuing Theory. *Queuing Theory and Telecommunications*, p. 305–383, 4 ago. 2005. Disponível em: <https://link.springer.com/chapter/10.1007/0-387-24066-7_5>. Acesso em: 14 jan. 2025.

MASEK, J. G. *et al.* Landsat 9: Empowering open science and applications through continuity. *Remote Sensing of Environment*, v. 248, 2020.

MEIRA, P. S. S. Weakly Supervised Semantic Segmentation With Class Balancing Strategies Applied to Remote Sensing Imagery. 2024. Acesso em: 21 abr. 2025.

MENG, H. *et al.* Classification of inland lake water quality levels based on Sentinel-2 images using convolutional neural networks and spatiotemporal variation and driving factors of algal bloom. *Ecological Informatics*, v. 80, p. 102549, 1 maio 2024. Acesso em: 22 abr. 2025.

MO, Y. *et al.* Review the state-of-the-art technologies of semantic segmentation based on deep learning. *Neurocomputing*, v. 493, p. 626–646, 7 jul. 2022. Acesso em: 14 jan. 2025.

NAIR, V.; HINTON, G. E. Rectified Linear Units Improve Restricted Boltzmann Machines. 2008. Acesso em: 20 jan. 2025.

NYLÉN, T.; CALLE, M.; GONZALES-INCA, C. Unveiling coastal change across the Arctic with full Landsat collections and data fusion. *Remote Sensing of Environment*, v. 322, p. 114696, 15 maio 2025. Acesso em: 22 abr. 2025.

OLOFSSON, P. *et al.* Good practices for estimating area and assessing accuracy of land change. *Remote Sensing of Environment*, v. 148, p. 42–57, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.rse.2014.02.015>>.

PAUL, R. K. *et al.* Image Compression Scheme based on Optimized K-means Clustering and Higher-Level Decomposed DWT. *Procedia Computer Science*, v. 235, p. 642–655, 1 jan. 2024. Acesso em: 23 jan. 2025.

PENG, X. *et al.* Application of K-Means method to pattern recognition in on-line cable partial discharge monitoring. 2013. Disponível em: <<https://doi.org/10.1109/TDEI.2013.6518945>>. Acesso em: 23 jan. 2025.

PONZONI, F. J.; SHIMABUKURO, Y. E. *Sensoriamento remoto no estudo da vegetação*. 2. ed. São Paulo: Oficina do Texto, 2009. v. 1.

PRECHELT, L. Early Stopping — But When? *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 7700 LECTURE NO, p. 53–67, 2012. Disponível em: <https://link.springer.com/chapter/10.1007/978-3-642-35289-8_5>. Acesso em: 20 jan. 2025.

RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. 2015, [S.l: s.n.], 2015.

ROSENBLATT, F. THE PERCEPTRON: A PROBABILISTIC MODEL FOR INFORMATION STORAGE AND ORGANIZATION IN THE BRAIN 1. *Psychological Review*, v. 65, n. 6, p. 19–27, 1958. Acesso em: 20 jan. 2025.

ROUSSEEUW, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, v. 20, n. C, p. 53–65, 1 nov. 1987. Acesso em: 15 jan. 2025.

RUDORFF, B. F. T. *et al.* Remote Sensing Images to Detect Soy Plantations in the Amazon Biome-The Soy Moratorium Initiative. *Sustainability*, v. 4, n. 5, p. 1074–1088, 2012.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. *Nature* 1986 323:6088, v. 323, n. 6088, p. 533–536, 1986. Disponível em: <<https://www.nature.com/articles/323533a0>>. Acesso em: 20 jan. 2025.

SCHMIDHUBER, J. *Deep Learning in neural networks: An overview. Neural Networks*. [S.l.: s.n.], 2015

SHELHAMER, E.; LONG, J.; DARRELL, T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 39, n. 4, p. 640–651, 1 abr. 2017. Acesso em: 14 jan. 2025.

SILVA, F. S. *et al.* DESENVOLVIMENTO DE MODELOS PARA A ESTIMATIVA DA PRODUTIVIDADE DO ALGODOEIRO A PARTIR DE SÉRIES TEMPORAIS DE IMAGENS DE SATÉLITE. . [S.l.]: UNICEP. Disponível em: <<https://repositorio.unicep.com.br/handle/123456789/53>>. Acesso em: 20 jan. 2025. , 20 dez. 2022

SIMIONATO, J.; BERTANI, G.; OSAKO, L. S. Identification of artisanal mining sites in the Amazon Rainforest using Geographic Object-Based Image Analysis (GEOBIA) and Data Mining techniques. *Remote Sensing Applications: Society and Environment*, v. 24, 2021.

SINGH, A.; YADAV, A.; RANA, A. K-means with Three different Distance Metrics. *International Journal of Computer Applications*, v. 67, n. 10, p. 13–17, 18 abr. 2013. Acesso em: 23 jan. 2025.

SONTER, L. J. *et al.* Mining drives extensive deforestation in the Brazilian Amazon. *Nature Communications*, 2017.

SONTER, L. J. *et al.* Processes of land use change in mining regions. *Journal of Cleaner Production*, v. 84, n. 1, p. 494–501, 2014.

SOUZA, C. M. *et al.* Reconstructing three decades of land use and land cover changes in brazilian biomes with landsat archive and earth engine. *Remote Sensing*, v. 12, n. 17, p. 2735, 1 set. 2020. Disponível em: <www.mdpi.com/journal/remotesensing>. Acesso em: 24 jan. 2021.

STRUDEL, R. *et al.* Segmenter: Transformer for Semantic Segmentation. *Proceedings of the IEEE International Conference on Computer Vision*, p. 7242–7252, 12 maio 2021. Disponível em: <<https://arxiv.org/abs/2105.05633v3>>. Acesso em: 15 jan. 2025.

SUN, YIFEI *et al.* Accurate mapping of rapeseed fields in the initial flowering stage using Sentinel-2 satellite images and convolutional neural networks. *Ecological Indicators*, v. 162, p. 112027, 1 maio 2024. Acesso em: 22 abr. 2025.

SUN, YUCHAO *et al.* Annual change in the distribution and landscape health of mangrove ecosystems in China from 2016 to 2023 with Sentinel imagery. *Global Ecology and Conservation*, v. 57, p. e03355, 1 jan. 2025. Acesso em: 22 abr. 2025.

SUWANDA, R.; SYAHPUTRA, Z.; ZAMZAMI, E. M. Analysis of Euclidean Distance and Manhattan Distance in the K-Means Algorithm for Variations Number of Centroid K. *Journal of Physics: Conference Series*, v. 1566, n. 1, 3 jul. 2020. Disponível em: <https://www.researchgate.net/publication/342675533_Analysis_of_Euclidean_Distance_and_Manhattan_Distance_in_the_K-Means_Algorithm_for_Variations_Number_of_Centroid_K>. Acesso em: 23 jan. 2025.

TASSO AZEVEDO. *MapBiomias Brasil*. Disponível em: <<https://brasil.mapbiomas.org/estatisticas/>>. Acesso em: 13 jan. 2025.

TELMER, K. *et al.* The source and fate of sediment and mercury in the Tapajós River, Pará, Brazilian Amazon: Ground- and space-based evidence. *Journal of Environmental Management*, v. 81, n. 2, p. 101–113, 1 out. 2006. Acesso em: 13 jan. 2025.

TOLAN, J. *et al.* Very high resolution canopy height maps from RGB imagery using self-supervised vision transformer and convolutional decoder trained on aerial lidar. *Remote Sensing of Environment*, v. 300, p. 113888, 1 jan. 2024. Acesso em: 22 abr. 2025.

TRAN, K. H. *et al.* A transformer-based model for detecting land surface phenology from the irregular harmonized Landsat and Sentinel-2 time series across the United States. *Remote Sensing of Environment*, v. 320, p. 114656, 1 abr. 2025. Acesso em: 22 abr. 2025.

URKIDI, L.; WALTER, M. Dimensions of environmental justice in anti-gold mining movements in Latin America. *Geoforum*, v. 42, n. 6, 2011.

VAN HOUDT, G.; MOSQUERA, C.; NÁPOLES, G. A review on the long short-term memory model. *Artificial Intelligence Review*, v. 53, n. 8, 2020.

VELÁSQUEZ RAMÍREZ, M. G. *et al.* Mercury in soils impacted by alluvial gold mining in the Peruvian Amazon. *Journal of Environmental Management*, v. 288, 2021.

WANG, B. *et al.* Bamboo classification based on GEDI, time-series Sentinel-2 images and whale-optimized, dual-channel DenseNet: A case study in Zhejiang province, China. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 209, p. 312–323, 1 mar. 2024. Acesso em: 22 abr. 2025.

WANG, C. *et al.* DLAU: A scalable deep learning accelerator unit on FPGA. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, v. 36, n. 3, 2017.

WANG, RAY; CHOWDHURY, T.; ORTIZ, A. C. Semantic segmentation framework for atoll satellite imagery: An in-depth exploration using UNet variants and Segmentation Gym. *Applied Computing and Geosciences*, v. 25, p. 100217, 1 fev. 2025. Acesso em: 22 abr. 2025.

WANG, RUIKUN *et al.* Transformers for Remote Sensing: A Systematic Review and Analysis. *Sensors* 2024, Vol. 24, Page 3495, v. 24, n. 11, p. 3495, 29 maio 2024. Disponível em: <<https://www.mdpi.com/1424-8220/24/11/3495/htm>>. Acesso em: 14 jan. 2025.

WANG, Y. *et al.* Entity-based image analysis: A new strategy to map rural settlements from Landsat images. *Remote Sensing of Environment*, v. 318, p. 114549, 1 mar. 2025. Acesso em: 22 abr. 2025.

WULDER, M. A. *et al.* Fifty years of Landsat science and impacts. *Remote Sensing of Environment*, v. 280, p. 113195, 1 out. 2022. Acesso em: 23 jan. 2025.

XIE, E. *et al.* SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers. *Advances in Neural Information Processing Systems*, v. 15, p. 12077–12090, 31 maio 2021. Disponível em: <<https://arxiv.org/abs/2105.15203v3>>. Acesso em: 15 jan. 2025.

XU, Z. *et al.* HRCNet: High-Resolution Context Extraction Network for Semantic Segmentation of Remote Sensing Images. *Remote Sensing 2021, Vol. 13, Page 71*, v. 13, n. 1, p. 71, 27 dez. 2020. Disponível em: <<https://www.mdpi.com/2072-4292/13/1/71/htm>>. Acesso em: 22 jan. 2025.

YANG, J. *et al.* Applicability analysis of weakly supervised semantic segmentation for identifying salinized soil boundaries. *Journal of Arid Environments*, v. 229, p. 105372, 1 ago. 2025. Acesso em: 22 abr. 2025.

YANG, Q. *et al.* Research on Improved U-net Based Remote Sensing Image Segmentation Algorithm. 22 ago. 2024. Disponível em: <<https://arxiv.org/abs/2408.12672v1>>. Acesso em: 22 jan. 2025.

YANG, Y. *et al.* Detecting the dynamics of vegetation disturbance and recovery in surface mining area via Landsat imagery and LandTrendr algorithm. *Journal of Cleaner Production*, v. 178, p. 353–362, 2018. Disponível em: <<https://doi.org/10.1016/j.jclepro.2018.01.050>>.

YANN LECUN, YOSHUA BENGIO, G. H. Deep learning (2015), Y. LeCun, Y. Bengio and G. Hinton. *Nature*, 2015.

YAO, C. *et al.* A real-time crop lodging recognition method for combine harvesters based on machine vision and modified DeepLab V3+. *Smart Agricultural Technology*, v. 11, p. 100926, 1 ago. 2025. Acesso em: 22 abr. 2025.

YE, S.; ZHU, Z.; CAO, G. Object-based continuous monitoring of land disturbances from dense Landsat time series. *Remote Sensing of Environment*, v. 287, p. 113462, 15 mar. 2023. Acesso em: 22 abr. 2025.

YU, Y. *et al.* Techniques and Challenges of Image Segmentation: A Review. *Electronics* 2023, Vol. 12, Page 1199, v. 12, n. 5, p. 1199, 2 mar. 2023. Disponível em: <<https://www.mdpi.com/2079-9292/12/5/1199/htm>>. Acesso em: 21 jan. 2025.

ZABALZA, J. *et al.* Novel segmented stacked autoencoder for effective dimensionality reduction and feature extraction in hyperspectral imaging. *Neurocomputing*, v. 185, 2016.

ZHANG, F. *et al.* A dual-path model merging CNN and RNN with attention mechanism for crop classification. *European Journal of Agronomy*, v. 159, p. 127273, 1 set. 2024. Acesso em: 22 abr. 2025.

ZHANG, Y. *et al.* Deep multimodal fusion for semantic image segmentation: A survey. *Image and Vision Computing*, v. 105, p. 104042, 1 jan. 2021. Acesso em: 14 jan. 2025.

ZHANG, Z.; LIU, Q.; WANG, Y. Road Extraction by Deep Residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, 2018.

ZHAO, Y. *et al.* Cloud Removal Using Multimodal GAN With Adversarial Consistency Loss. *IEEE Geoscience and Remote Sensing Letters*, 2021.

ZHENG, F. *et al.* A Lightweight Dual-Branch Swin Transformer for Remote Sensing Scene Classification. *Remote Sensing* 2023, Vol. 15, Page 2865, v. 15, n. 11, p. 2865, 31 maio 2023. Disponível em: <<https://www.mdpi.com/2072-4292/15/11/2865/htm>>. Acesso em: 22 jan. 2025.

ZHOU, T.; RUAN, S.; CANU, S. A review: Deep learning for medical image segmentation using multi-modality fusion. *Array*, v. 3–4, p. 100004, 1 set. 2019. Acesso em: 14 jan. 2025.

ZHOU, Z. *et al.* Unet++: A nested u-net architecture for medical image segmentation. 2018, [S.l: s.n.], 2018.

ZHU, X. X. *et al.* *Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources*. *IEEE Geoscience and Remote Sensing Magazine*. [S.l: s.n.], , 2017

