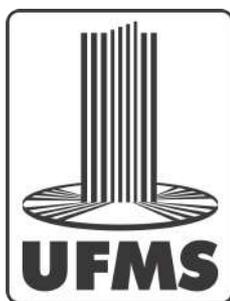


Paulo Augusto Arantes Vilela

**Deteccção de Desmatamento no Estado de  
Mato Grosso do Sul utilizando Segmentação  
Semântica em imagens bi-temporais dos  
satélites Landsat 8 e Sentinel 2**



Campo Grande - MS

2024



Paulo Augusto Arantes Vilela

# **Deteccção de Desmatamento no Estado de Mato Grosso do Sul utilizando Segmentação Semântica em imagens bi-temporais dos satélites Landsat 8 e Sentinel 2**

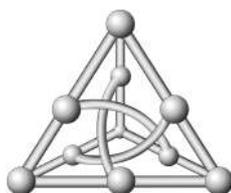
Dissertação apresentada ao Programa de Pós-Graduação da Faculdade de Computação (FACOM) da Universidade Federal de Mato Grosso do Sul (UFMS), como parte dos requisitos necessários à obtenção do título de Mestre em Ciência da Computação.

Universidade Federal de Mato Grosso do Sul – UFMS

Faculdade de Computação

Mestrado em Ciência da Computação

Orientador: Prof. Dr. Edson Takashi Matsubara



Campo Grande - MS

2024

# Resumo

A supressão da vegetação nativa sem autorização no Estado de Mato Grosso do Sul tem ameaçado a conservação dos biomas locais como o Cerrado e a Mata Atlântica. Neste contexto, é relevante salientar que a Constituição Brasileira de 1988 fixou como dever do Estado e da Coletividade a defesa do meio ambiente, atribuindo ao Ministério Público funções de atuação para a proteção ambiental. A exemplo, o Ministério Público do Estado de Mato Grosso do Sul, por meio do Programa DNA Ambiental, têm monitorado e identificado pontos de desmatamento sem autorização ambiental, em seu Núcleo de Geotecnologias (NUGEO), promovendo ações e providências para a punição e reparação de eventuais danos causados. Atualmente, este mapeamento tem sido realizado através da análise visual, não automatizada, de imagens de diversos satélites, especialmente dos satélites Landsat-8 e Sentinel-2, demandando consideráveis horas de trabalho e mão de obra especializada. Entretanto, com a evolução de técnicas de Deep Learning, novos algoritmos podem ser capazes de automatizar o processo de análise multitemporal de imagens de satélites, promovendo agilidade, ganho de eficiência e possibilitando a alocação de recursos humanos para outros serviços. Este trabalho tem por objetivo apresentar uma proposta de automatização do processo de identificação de desmatamentos utilizando as redes neurais profundas DeepLabv3+, U-Net e Multi-Scale Attention for Semantic Segmentation, para segmentação semântica em imagens bi-temporais de satélite, disponibilizando um modelo de inteligência artificial treinado, com a capacidade de mapear áreas desmatadas em qualquer cena dos biomas Cerrado e Mata Atlântica. Para tanto, buscou-se um conjunto de dados apropriado para essa finalidade. O primeiro dataset foi gerado a partir de arquivo shapefile com polígonos de áreas desmatadas pré-determinadas, associado ao recorte manual de duas cenas do satélite Landsat-8, resultando em um conjunto de dados com poucas amostras e muito ruído. O segundo dataset foi obtido a partir de um novo arquivo shapefile de desmatamentos e cenas do satélite Sentinel-2, cujo download e recortes foram realizados de forma automatizada utilizando-se um framework desenvolvido com APIs e infraestrutura em nuvem do projeto Planetary Computer, associadas a aplicação de uma curadoria especializada, produzindo um conjunto com maior número de amostras e ruído reduzido, com o qual obteve-se os melhores resultados nos treinamentos preliminares utilizando-se a rede neural U-net, adotada como baseline, sendo então escolhido para o treinamento das demais redes neurais artificiais avaliadas neste estudo. Ainda, foi realizado experimento para mitigar o problema do balanceamento dos dados, empregando-se diferentes funções de perda. Assim sendo, as principais contribuições desta dissertação são: (1) um conjunto de dados de desmatamento público rotulado e com curadoria; (2) uma avaliação experimental usando as redes neurais U-Net, DeepLabv3+ e Multi-Scale Attention para segmentação semântica em imagens bi-temporais do satélite Sentinel 2; e (3) uma avaliação experimental das funções de perda Binary-Cross-Entropy, Weighted Binary-Cross-Entropy e Focal Loss. Os resultados mais favoráveis foram obtidos com a arquitetura Multi-Scale Attention for Semantic Segmentation, com a função de perda Weighted Binary Cross-Entropy. Por fim, esse modelo foi disponibilizado ao MPMS, que procederá a sua implantação e implementação de procedimentos para validação dos resultados, bem como aprimoramento da amostra e evolução do modelo.

**Palavras-chaves:** Detecção de Desmatamento; Cerrado; Landsat; Sentinel; Segmentação Semântica; Deep Learning.

# Abstract

The unauthorized suppression of native vegetation in the state of Mato Grosso do Sul has threatened the conservation of local biomes such as the Cerrado and the Atlantic Forest. In this context, it is relevant to emphasize that the Brazilian Constitution of 1988 established as duty of the State and Society the defense of a balanced environment, assigning to the Public Prosecutor's Office functions of action for environmental protection. For example, the Public Prosecutor's Office of the State of Mato Grosso do Sul, through the DNA Ambiental Program, has been monitoring and identifying points of deforestation without environmental authorization in its Geotechnologies Center (NUGEO), promoting actions and measures for the punishment and reparation of any damages caused. Currently, this mapping has been carried out through visual, non-automated analysis of images from various satellites, especially Landsat-8 and Sentinel-2 satellites, requiring considerable hours of work and specialized labor. However, with the evolution of Deep Learning techniques, new algorithms may be able to automate the multitemporal analysis process of satellite images, promoting agility, efficiency gains, and enabling the allocation of human resources to other services. This work aimed to present a proposal for automating the process of identifying deforestation using the deep neural networks DeepLabv3+, U-Net, and Multi-Scale Attention, for semantic segmentation in bi-temporal images of satellite, providing a trained artificial intelligence model capable of mapping deforested areas in any scene of the Cerrado and Atlantic Forest biomes. For this purpose, an appropriate dataset was sought. The first dataset was generated from a shapefile with predefined deforested areas polygons, associated with the manual clipping of two Landsat-8 satellite scenes, resulting in a dataset with few samples and much noise. The second dataset was obtained from a new deforestation shapefile and Sentinel-2 satellite scenes, with downloads and clippings performed automatically using a framework developed with APIs and cloud infrastructure from the Planetary Computer project, associated with specialized curation, producing a dataset with a larger number of samples and reduced noise. Notably, this dataset yielded the best results in preliminary training using the U-net neural network, established as the baseline, and subsequently adopted for training other evaluated artificial neural networks in this study. Additionally, an experiment was conducted to address data imbalance issues, employing various loss functions. Therefore, the main contributions of this dissertation are: (1) a labeled and curated public deforestation dataset; (2) an experimental evaluation using the U-Net, DeepLabv3+, and Multi-Scale Attention neural networks for semantic segmentation in bi-temporal Sentinel-2 satellite images; and (3) an experimental evaluation of the loss functions Binary-Cross-Entropy, Weighted Binary-Cross-Entropy, and Focal Loss. The most favorable outcomes were achieved with the Multi-Scale Attention for Semantic Segmentation architecture, utilizing the Weighted Binary Cross-Entropy loss function. Finally, the most effective model has been made accessible to the MPMS for deployment, implementation of validation procedures, and further refinement of samples and model evolution.

**Key-words:** Deforestation mapping; Cerrado; Landsat; Sentinel; Semantic Segmentation; Deep Learning.



# Lista de ilustrações

Figura 1 – Ilustração do processo de análise . . . . .	15
Figura 2 – Ilustração dos dados disponibilizados inicialmente. . . . .	16
Figura 3 – Exemplo de segmentação . . . . .	23
Figura 4 – Arquitetura da U-Net . . . . .	26
Figura 5 – Ilustração de uma convolução transposta . . . . .	26
Figura 6 – Ilustração do modelo Multi-Scale Attention . . . . .	28
Figura 7 – Limite dos municípios no Estado de Mato Grosso do Sul. . . . .	38
Figura 8 – Exemplo de polígonos de desmatamento disponibilizados no arquivo Shapefile. . . . .	38
Figura 9 – Exemplo de área com desmatamento e sem desmatamento. . . . .	39
Figura 10 – Metodologia inicial - Processamento das cenas do satélite Landsat-8. . . . .	40
Figura 11 – Área de estudo relativa ao segundo conjunto de dados. . . . .	41
Figura 12 – Nova metodologia de preparação do dataset - Sentinel 2. . . . .	42
Figura 13 – Exemplo de recorte removido da amostra por conter possível desmate no canto superior direito não identificado. . . . .	42
Figura 14 – Exemplo de recorte removido da amostra por conter falha na geração do recorte. . . . .	42
Figura 15 – Exemplo de recorte removido da amostra por conter quantidade excessiva de nuvens. . . . .	43
Figura 16 – Exemplo de recorte mantido na amostra por ter sido considerado ideal. . . . .	43
Figura 17 – Metodologia de treinamento. . . . .	43
Figura 18 – Histograma - Quantidade de exemplos por percentual de cobertura de área desmatada. . . . .	44
Figura 19 – Boxplot - Distribuição uniforme das porções de área desmatada para todos os conjuntos. . . . .	44
Figura 20 – Conjunto de dados <i>Sentinel 2</i> . . . . .	45
Figura 21 – Representação visual do processo de mapeamento do desmatamento.. . . .	46
Figura 22 – Exemplo de resultado visualmente promissor utilizando o conjunto de dados Landsat-8. . . . .	49
Figura 23 – Exemplo de resultado do conjunto de dados Landsat-8 com defeito. . . . .	50
Figura 24 – Curva de perda durante o treinamento da arquitetura <i>Multi-Scale Attention for Semantic Segmentation</i> . . . . .	51
Figura 25 – Curva ROC das arquiteturas U-net, DeepLabv3+ e Multi-Scale Attention. . . . .	52

Figura 26 – Curva ROC da arquitetura Multi-scale attention com diferentes funções de perda. . . . .	53
Figura 27 – Exemplo de resultado 1 - Arquitetura Multi-Scale Attention com a função de perda Weighted Binary Cross-Entropy. . . . .	55
Figura 28 – Exemplo de resultado 2 - Arquitetura Multi-Scale Attention com a função de perda Weighted Binary Cross-Entropy. . . . .	55
Figura 29 – Exemplo de resultado 3 - Arquitetura Multi-Scale Attention com a função de perda Weighted Binary Cross-Entropy. . . . .	56
Figura 30 – Exemplo de resultado 4 - Arquitetura Multi-Scale Attention com a função de perda Weighted Binary Cross-Entropy. . . . .	56

# Lista de tabelas

Tabela 1 – Comparação entre resultados recentes para detecção de desmatamento na Amazônia utilizando aprendizado profundo. . . . .	32
Tabela 2 – Comparação das abordagens utilizadas para detecção de desmatamento no bioma Cerrado. . . . .	35
Tabela 3 – Comparação entre resultados dos conjuntos de dados Lansat-8 e Sentinel 2. . . . .	49
Tabela 4 – Comparação entre resultados das arquiteturas U-net, DeepLabv3+ and Multi-Scale Attention segmentados por classe utilizando Binary Cross-Entropy como função de perda. . . . .	51
Tabela 5 – Impacto da utilização de diferentes funções de perda utilizando a arquitetura Multi-scale Attention. Os valores em negrito representam o maior valor obtido para a classe desmatada. . . . .	53
Tabela 6 – Resultado do modelo Multi-Scale Attention com a função de perda Weighted Binary Cross-Entropy, aplicando Regressão Isotônica. . . . .	54



# Lista de abreviaturas e siglas

API	<i>Application Programming Interface</i>
ASD	<i>Average Soft Dice</i>
ASPP	<i>Atrous Spatial Pyramid Pooling</i>
CAOMA	<i>Centro de Apoio Operacional das Promotorias da Justiça do Meio Ambiente, da Habitação e Urbanismo e do Patrimônio Histórico e Cultural</i>
DLCD	<i>DeepLab-based Change Detection</i>
EF	<i>Early Fusion</i>
EVI	<i>Enhanced Vegetation Index</i>
FCN	<i>Fully Convolutional Neural Networks</i>
FN	<i>Falsos Negativos</i>
FP	<i>Falsos Positivos</i>
Ibama	<i>Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis</i>
IoU	<i>Intersection over Union</i>
LF	<i>Late Fusion</i>
MLP	<i>Multilayer Perceptron</i>
MPMS	<i>Ministério Público do Estado de Mato Grosso do Sul</i>
NDVI	<i>Normalized Difference Vegetation Index</i>
NUGEO	<i>Núcleo de Geotecnologias do MPMS</i>
RF	<i>Random Forest</i>
SVM	<i>Support Vector Machine</i>
SWIR	<i>Short-wave Infrared</i>
S-CNN	<i>Siamese Convolutional Neural Network</i>
VP	<i>Verdadeiros Positivos</i>

WFL      *Weighted Focal Loss*  
WCE      *Weighted Cross-Entropy*

# Sumário

<b>1</b>	<b>Introdução</b>	<b>13</b>
1.1	Hipótese	16
1.2	Objetivos	16
1.3	Organização	17
<b>2</b>	<b>Fundamentação</b>	<b>19</b>
2.1	Desmatamento	19
2.1.1	Cerrado	20
2.1.2	Mata Atlântica	20
2.2	Sensoriamento remoto	21
2.2.1	Geoprocessamento	21
2.3	Ministério Público	22
2.4	Aprendizado de máquina	23
2.4.1	Segmentação semântica	23
2.4.2	Segmentação semântica para detecção de mudanças	24
2.4.3	U-NET	25
2.4.4	<i>DeepLab-based Change Detection (DLCD)</i>	27
2.4.5	<i>Hierarchical Multi-Scale Attention for Semantic Segmentation</i>	28
2.4.6	Métricas de avaliação	28
<b>3</b>	<b>Trabalhos Relacionados</b>	<b>31</b>
3.1	Amazônia	31
3.2	Cerrado e Mata Atlântica	34
<b>4</b>	<b>Metodologia</b>	<b>37</b>
4.1	Primeiro conjunto de dados – Landsat 8	37
4.2	Segundo conjunto de dados – Sentinel 2	40
4.3	Experimentos	44
<b>5</b>	<b>Resultados e Discussão</b>	<b>49</b>
5.1	Experimento 1 - Comparação dos conjuntos de dados	49
5.2	Experimento 2 - Arquitetura da rede	50
5.3	Experimento 3 - Balanceamento do conjunto de dados	52
5.4	Experimento 4 - Calibração do classificador	54
5.5	Inspeção visual e análise de resultados	54
<b>6</b>	<b>Considerações Finais</b>	<b>57</b>

**Referências** . . . . . 59

# 1 Introdução

O Estado de Mato Grosso do Sul se caracteriza como um importante produtor de soja, milho, açúcar, celulose e carne bovina no cenário brasileiro (FAGUNDES *et al.*, 2017). Situado na região Centro-Oeste do país, em meio aos biomas Cerrado, Mata Atlântica e Pantanal (IBGE, 2019), se beneficia com solos férteis e extensas superfícies planas, cujo clima, predominantemente tropical e subtropical, favorece a produção agropecuária, típica da região que investe em tecnologias voltadas para campo, visando aumentar a produção e exportação de seus produtos (TETILA *et al.*, 2020).

No entanto, a expansão das atividades agrícolas e pecuárias, frequentemente marcada pela conversão de áreas com vegetação nativa em áreas de monocultura e criação de animais, pelo uso intensivo de agrotóxicos, fertilizantes e corretivos agrícolas, bem como pelo pisoteio excessivo do gado (CUNHA *et al.*, 2008), têm ameaçado a conservação de ecossistemas, como o Cerrado (RUMBLE *et al.*, 2019; VACCHIANO *et al.*, 2018) e a Mata Atlântica (MYERS *et al.*, 2000; PAIM, 2020).

O Cerrado é caracterizado pela sua ampla endemia de espécies e alta diversidade biológica (SILVA *et al.*, 2006). Entretanto, a perda da vegetação, em decorrência das atividades antrópicas, tem resultado na redução da biodiversidade e no aumento das espécies ameaçadas de extinção neste bioma (RODRIGUES *et al.*, 2020).

A Mata Atlântica, reconhecida como um dos biomas mais vulneráveis do Brasil, possui apenas 12,4% de sua cobertura original em território nacional, dos quais 80% estão localizados em áreas privadas, conforme levantamentos realizados pela [Fundação SOS Mata Atlântica \(2019\)](#), podendo ser alvos de intervenções. A exemplo, somente no período de janeiro a outubro de 2022, o desmatamento acumulado nesse bioma foi de 48.660 hectares ([Fundação SOS Mata Atlântica, 2022](#)).

O vertiginoso processo de expansão agropecuária no Estado de Mato Grosso do Sul tem suscitado o desmatamento em áreas protegidas por lei e diferentes graus de degradação ambiental (GONÇALVES *et al.*, 2021). Neste contexto, é relevante salientar que a Constituição Federal Brasileira de 1988, atribuiu como deveres do Estado e da Coletividade a defesa e a preservação do meio ambiente ecologicamente equilibrado, para as presentes e futuras gerações (BRASIL, 1988). Ainda, concedeu ao Ministério Público funções de atuação em sua preservação, como a promoção de Inquérito Civil e Ação Civil Pública para a proteção ambiental e de outros interesses difusos e coletivos. Assim sendo, o órgão ministerial tem atuado efetivamente na esfera cível com intuito de resguardar o meio ambiente e exigir, caso necessário e conforme a legislação vigente, a reparação de eventuais danos causados (OLLIVEIRA; MEDEIROS; DIETRICH, 2018).

A exemplo, o Ministério Público do Estado de Mato Grosso do Sul (MPMS) criou o Programa de Detecção de Desmatamento Ilegal de Vegetação Nativa (Programa DNA Ambiental), com intuito de combater a supressão ilegal de áreas nativas através do monitoramento periódico dos desmatamentos, realizado por meio da análise bi-temporal de imagens de satélites por seu Núcleo de Geotecnologias (NUGEO). Esse identifica polígonos de desmatamento sem autorização ambiental e emite pareceres aos órgãos ambientais e às Promotorias de Justiça para a tomada de providências quanto a responsabilização dos infratores e reparação dos danos constatados (GONÇALVES et al., 2021).

Por meio do Programa DNA Ambiental foram identificados, apenas entre os meses de janeiro e fevereiro de 2020, 5.787,76 hectares de desmatamento em vegetação nativa no Estado de Mato Grosso do Sul, dos quais 4.808,4 hectares dispunham de autorizações ambientais para supressão vegetal, restando 979,36 hectares de desmatamento possivelmente ilegal (MPMS, 2020).

Atualmente os técnicos do MPMS realizam a análise visual, não automatizada, de imagens de diversos satélites em formato GeoTIFF, especialmente dos satélites Landsat-8 e Sentinel 2, comparando cenas de períodos distintos, em busca de alterações na cobertura vegetal com indicações de possíveis desmatamentos, utilizando a ferramenta Quantum GIS (QGIS). Essa tarefa é onerosa e demanda tempo e mão de obra especializada. Dessa forma, existem diversas vantagens ao considerar a automatização deste trabalho manual, como o ganho de eficiência e a possibilidade de alocação de recursos humanos para outros serviços, podendo ser empregadas técnicas de aprendizado de máquina com essa finalidade.

Na área da computação, o problema de detecção de desmatamento é considerado como integrado nas categorias que buscam solução em técnicas de detecção de mudança, aplicada ao monitoramento de florestas, em recursos ambientais (SHI et al., 2020). Neste contexto, Bem et al. (2020) avaliaram a utilização de técnicas de aprendizado profundo para a atividade de detecção de desmatamento na Amazônia e constataram significantes melhorias nos resultados obtidos em comparação com os resultados de algoritmos clássicos de aprendizado de máquina.

Estudos recentes apontam para resultados promissores com a utilização de modelos de redes neurais profundas para detecção de desmatamento no mundo (SOLÓRZANO et al., 2023; ZHAO et al., 2022; ISAIENKOV et al., 2021; LEE et al., 2020; KATARKI et al., 2019; KHAN et al., 2017), e no Brasil com ênfase para o bioma Amazônico (ALSHEHRI; OUADOU; SCOTT, 2024; PARK et al., 2023; TORRES et al., 2021; NAZAROVA; MARTIN; GIULIANI, 2020; ANDRADE et al., 2020; MARETTO et al., 2020; ORTEGA et al., 2019). Já para os biomas Cerrado e Mata Atlântica, a quantidade de publicações é consideravelmente menor, ainda que com bons resultados (MATOSAK et al., 2022; ADARME et al., 2020; PARENTE et al., 2019) e, considerando a relatividade das métricas utilizadas para avaliação, apenas estes dados não seriam suficientes para declarar o assunto como

esclarecido. Ademais, não foi encontrado na literatura qualquer proposta de aplicação desta tecnologia especificamente no meio onde ela é necessária, possibilitando, mesmo que indiretamente, a contribuição efetiva para a preservação do ambiente.

Nestas circunstâncias, o presente trabalho apresentou uma proposta de automação do processo de identificação de desmatamentos (Figura 1) utilizando redes neurais profundas, treinadas a partir de um conjunto de dados produzido utilizando os artefatos disponibilizados pela equipe do MPMS para este projeto.

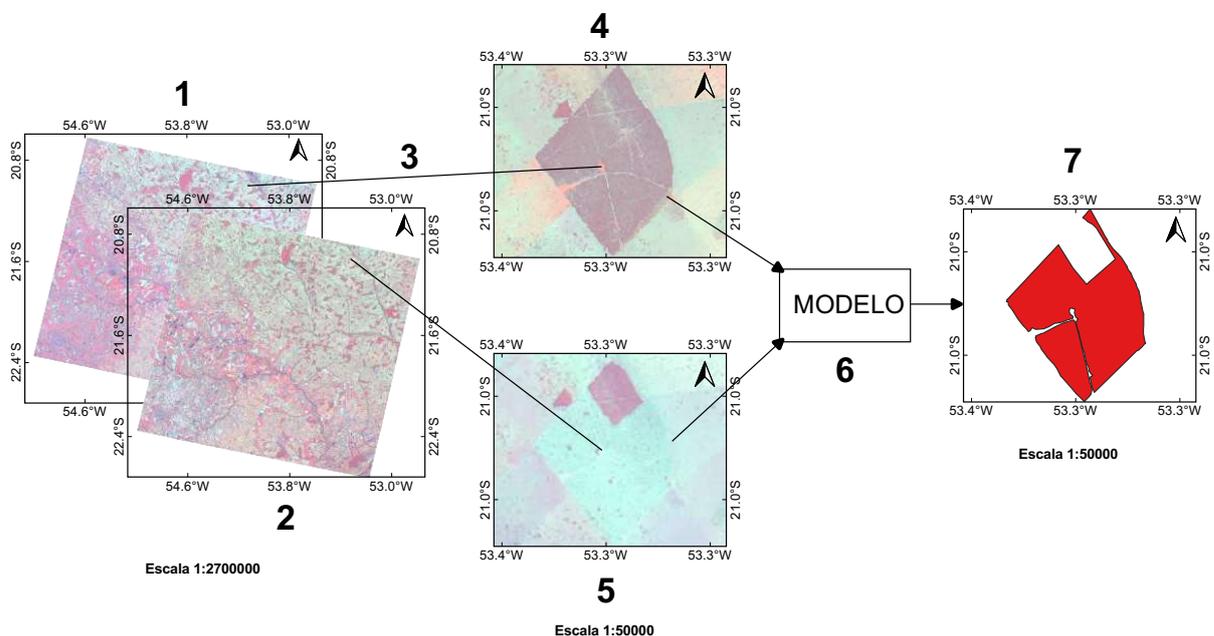


Figura 1 – Ilustração do processo de análise: 1 - Cena do satélite Landsat-8 de 2015. 2 - Cena do satélite Landsat-8 de 2018. 3 - Pré-processamento. 4 - Exemplo de recorte da cena de 2015. 5 - Exemplo de recorte da cena de 2018. 6 - Modelo de aprendizado de máquina. 7 - Saída esperada do modelo.

Inicialmente foram fornecidas duas imagens do satélite Landsat-8 de uma mesma cena, sobreposta a áreas dos biomas Mata Atlântica e Cerrado, referentes aos meses de agosto de 2015 e junho de 2018, contendo também um arquivo *shapefile* com polígonos onde o desmatamento foi identificado nesta área durante o período entre as duas datas (Figura 2). Em um segundo momento, foram fornecidos novos arquivos no formato *shapefile* contendo polígonos com a delimitação de áreas desmatadas em todo o território do Estado de Mato Grosso do Sul, possibilitando a geração automatizada de um novo conjunto de dados, com maior quantidade de recortes, utilizando imagens do satélite Sentinel 2.

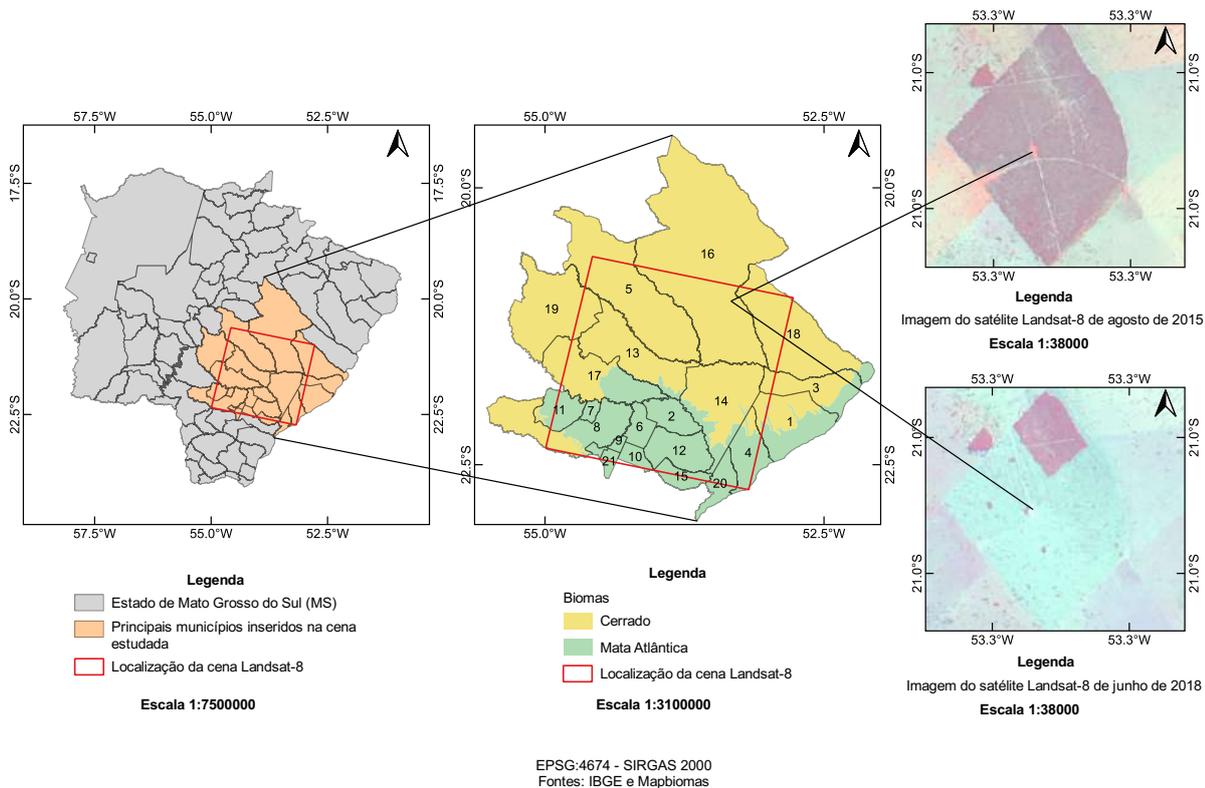


Figura 2 – Ilustração dos dados disponibilizados inicialmente.

## 1.1 Hipótese

Com a evolução de técnicas de Deep Learning, as arquiteturas DeepLabv3+, U-Net e Multi-Scale Attention for Semantic Segmentation podem ser capazes de realizar análise bi-temporal de imagens de satélites para identificação de supressão de vegetação nativa com assertividade próxima à das análises realizadas por técnicos do NUGEO, especialmente para áreas nos biomas Cerrado e Mata Atlântica.

## 1.2 Objetivos

O objetivo geral deste trabalho foi analisar a efetividade de diferentes arquiteturas de redes neurais profundas (DeepLabv3+, U-Net e Multi-Scale Attention for Semantic Segmentation) para segmentação semântica em imagens bi-temporais de satélite, para fins de detecção de desmatamentos.

Como objetivos específicos este trabalho apresentou algumas soluções para os seguintes questionamentos:

1. Considerando os artefatos disponibilizados, o custo de pré-processamento, de curadoria e a qualidade dos conjuntos de dados resultantes para serem utilizados no

treinamento de redes neurais artificiais e demais experimentos, qual a metodologia mais adequada para produção da amostra?

2. Considerando o desbalanceamento do conjunto de dados disponibilizado, tendo em vista que as amostras apresentam um considerável número de pixels classificados como sem desmate, qual técnica poderá ser utilizada para viabilizar o aprendizado nestes conjuntos desbalanceados?
3. Qual função de perda é mais adequada para a atividade de detecção de desmatamento?

## 1.3 Organização

Este trabalho está organizado da seguinte forma: o Capítulo 2 é responsável pela fundamentação teórica, com a apresentação dos principais conceitos e tecnologias utilizadas; o Capítulo 3 sintetiza os resultados das recentes publicações relacionadas à atividade de detecção de desmatamento utilizando aprendizado de máquina; o Capítulo 4 dispõe sobre a metodologia e caracterização do conjunto de dados; o Capítulo 5 apresenta os resultados dos experimentos realizados e, por último, no Capítulo 6 encontram-se as considerações finais.



## 2 Fundamentação

Neste capítulo temos a exposição dos principais conceitos relacionados a atividade de detecção de desmatamento, com uma abordagem objetiva sobre as tecnologias utilizadas.

### 2.1 Desmatamento

A expansão da agricultura e da pecuária, latentes no Estado de Mato Grosso do Sul, associadas ao crescimento desordenado das áreas urbanas (CUNHA et al., 2008), têm propiciado um cenário de intensas transformações e problemas ambientais, exercendo forte ameaça aos biomas locais, como Cerrado (RUMBLE et al., 2019; VACCHIANO et al., 2018) e a Mata Atlântica (MYERS et al., 2000; PAIM, 2020).

A alteração da cobertura do solo, como resultado das atividades humanas, especialmente as agropecuárias, marcadas pelo desmatamento de áreas com vegetação nativa, têm desencadeado processos erosivos, assoreamento de rios, redução da qualidade das águas, perda de biodiversidade e reflexos sobre os biomas brasileiros (CUNHA et al., 2008).

Não apenas em Mato Grosso do Sul, ou nos demais Estados brasileiros, o desmatamento tem causado prejuízos ambientais em todo o planeta, tornando-se tema frequente de discussões no cenário internacional, com ênfase, especialmente, nas quantidades de carbono sequestradas pelas áreas de florestas nativas (GONÇALVES et al., 2021).

As florestas abrigam a maior biodiversidade terrestre, proporcionando habitats para 68% das espécies de mamíferos, 75% das espécies de aves e 80% das espécies de anfíbios (VIE; HILTON-TAYLOR; STUART, 2009). Todavia, apesar de um empenho global na proteção do meio ambiente, a quantidade total de florestas tem reduzido.

No período de 2010 a 2015, constatou-se redução a uma taxa de 5,5 milhões de ha por ano em área das florestas tropicais (KEENAN et al., 2015). Já no período de 2010 e 2020, só a América do Sul apresentou uma perda relativa de florestas de 2,6 milhões de hectares por ano (FAO, 2020). O Brasil, entre os anos de 2010 e 2015, ocupou a primeira posição no ranking de perda líquida de área florestal (FAO, 2016).

No que diz respeito ao bioma Mata Atlântica, consoante aos levantamentos realizados pela a [Fundação SOS Mata Atlântica \(2019\)](#), restam apenas 12,4% de sua cobertura original em território nacional, e desses remanescentes, 80% estão em áreas privadas. Quanto ao Cerrado, constata-se apenas 38% de sua cobertura natural em Mato Grosso do Sul, consoante ao Plano de Ação para Prevenção e Controle do Desmatamento e das Queimadas no Cerrado (BRASIL, 2018).

Neste cenário, insurge-se os avanços das tecnologias de sensoriamento remoto e dos sistemas de informações geográficas, essenciais para o monitoramento da cobertura vegetal (FERREIRA; FERREIRA; FERREIRA, 2008). A disponibilidade de dados de sensoriamento remoto permitiu, tanto em nível local, como global, a obtenção de informações espacialmente explícitas sobre a variação na cobertura florestal (ACHARD et al., 2006).

### 2.1.1 Cerrado

Na região de Mato Grosso do Sul, tal como em toda região Centro-Oeste do Brasil, a vegetação predominante é o Cerrado, cobrindo 61% do território do Estado (IMASUL, 2018). Este bioma se destaca por sua elevada biodiversidade e ampla endemia de espécies, sendo considerado um dos 25 *hotspots* de biodiversidade no mundo, contudo, tem sofrido consideravelmente com a perda da vegetação, decorrente da expansão das atividades agrícolas e pecuárias (RUMBLE et al., 2019; VACCHIANO et al., 2018).

Caracterizado como o segundo maior bioma brasileiro, o Cerrado é descrito como a savana mais rica do mundo, abrigando mais de 11.000 espécies de flora nativa (SILVA; BATES, 2002). Destas, 4.400 espécies de plantas são endêmicas (MYERS et al., 2000). Atualmente, evidenciam-se como principais ameaças a este bioma o desmatamento, o fogo e as gramíneas invasoras, que estão centradas na expansão da agricultura e pecuária (CUNHA et al., 2008).

### 2.1.2 Mata Atlântica

Presente em menores proporções do território nacional, destaca-se o Bioma Mata Atlântica, que ocorre em 17 Estados brasileiros e 16 capitais (Fundação SOS Mata Atlântica, 2019). Assim como o bioma Cerrado, a Mata Atlântica é considerada um *hotspot* de biodiversidade, abrigando mais de 1.500 espécies vegetais endêmicas (MYERS et al., 2000).

Este bioma abrange grandes centros urbanos e, por consequência, apresenta altos índices de degradação florestal decorrente das atividades humanas, como o desenvolvimento da pecuária e expansão da fronteira agrícola (COSTA; GOMES; ALMEIDA, 2014).

Caracterizado como um dos biomas mais ameaçados do planeta, ainda que protegido por leis ambientais específicas. De acordo com o Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (Ibama), dentre as 633 espécies de animais ameaçadas de extinção no Brasil, 383 ocorrem nesse bioma.

## 2.2 Sensoriamento remoto

O advento do termo sensoriamento remoto ocorreu na década de 60 (JENSEN, 2009), referindo-se ao uso de sensores capazes de “Captar a energia eletromagnética total ou parcialmente refletida ou emitida, por objetos e feições sob a superfície terrestre” (PEDRASSOLI, 2016).

O desenvolvimento desta tecnologia de coleta automática de dados para o levantamento da superfície terrestre está associado à revolução dos mecanismos de observação do planeta em uma escala global e periódica, tal como à velocidade da monitoração de alterações das feições terrestres e fenômenos dinâmicos (MENESES; ALMEIDA, 2012). Para Liu (2006), esta tecnologia evoluiu muito desde o lançamento do primeiro satélite, em 1957, sendo utilizada atualmente para diversas finalidades:

As aplicações atuais da Ciência de Sensoriamento Remoto via satélite abrangem monitoramento dinâmico de usos de solo e evoluções de biodiversidade, recursos naturais, precipitação, clima, manejo dos recursos hídricos da superfície, exploração dos recursos hídricos no subsolo, inventário e monitoramento de usos do solo, das pastagens e das coberturas vegetais, monitoramento da produção primária das florestas e pastagens, monitoramento das ocorrências dos eventos catastróficos, tais como geadas, enchentes, secas, doenças e pragas das culturas, monitoramento dos focos de queimadas, mapeamento topográfico, monitoramento da degradação e erosão de solos, estudo das mudanças climáticas, e outros. (LIU, 2006)

Dentre os principais satélites de monitoramento da superfície terrestre utilizados no sensoriamento remoto, destacam-se os satélites Landsat-8 e Sentinel-2.

O primeiro dispõe de imagens gratuitas, ortorretificadas, disponibilizadas em cenas contendo 185 por 180 Km com uma resolução espacial multiespectral de 30 metros por pixel e periodicidade de revisita de 16 dias (U.S. Geological Survey, 2019). Ressalta-se que ao realizar a fusão de bandas do satélite Landsat-8 com a Pancromática, a resolução espacial multiespectral passa a ser de 15 metros.

O segundo possui características semelhantes ao satélite Landsat-8, porém, suas imagens são disponibilizadas em cenas contendo 290 por 290 km com uma resolução espacial multiespectral de 10 metros por pixel com a periodicidade de revisita de 5 dias (European Space Agency, 2012).

### 2.2.1 Geoprocessamento

Para Silva, Monteiro e Pamboukian (2016), “geoprocessamento é o processo que relaciona localizações geográficas com diversos tipos de informações”. Um sistema de informações geográficas (SIG) pode ser definido como um conjunto de hardware, software

e informações demográficas com o propósito de viabilizar a utilização integrada de todas as formas de informações referenciadas geograficamente (PARANHOS et al., 2021).

Como um SIG é desenvolvido para funcionar com uma representação plana da terra, que naturalmente possui uma superfície irregular, algumas técnicas e padronizações precisam ser definidas de forma a garantir a correspondência entre as coordenadas independente da forma de projeção, tendo sido criado para esta finalidade os Sistemas de Projeção e os Sistemas de Referência de Coordenadas (SILVA; MONTEIRO; PAMBOUKIAN, 2016).

## 2.3 Ministério Público

O direito ao meio ambiente ecologicamente equilibrado, essencial à sadia qualidade de vida, é amparado pela Constituição Federal Brasileira, a qual atribuiu ao Estado e a Coletividade, o dever de defendê-lo e preservá-lo para as presentes e futuras gerações (BRASIL, 1988). Dentre os entes do Poder Público que atuam em sua proteção, destaca-se o Ministério Público, instituição permanente e essencial a função jurisdicional do Estado, a qual compete atuar por meio de Inquéritos Cíveis e as Ações Cíveis Públicas para a proteção ambiental e de outros interesses difusos e coletivos (BRASIL, 1988).

Diante de tal incumbência, o órgão tem utilizado as mais diversas ferramentas para otimizar a elaboração de diagnósticos sobre danos ambientais, com intuito de avaliar o cumprimento da legislação ambiental. Para tanto, ferramentas de geotecnologia vêm ganhando destaque em sua atuação, servindo de auxílio no planejamento de diligências em campo, planejamento de estratégias de conservação e preservação dos recursos naturais, bem como no monitoramento de aspectos ambientais (SANTOS; VALERIANO, 2003; RASLAN et al., 2009; MENDONÇA et al., 2011; TANCREDI et al., 2012).

A exemplo, no Ministério Público do Piauí, a punição dos responsáveis por danos ambientais tornou-se mais célere, em decorrência da aplicação de geotecnologias, contribuindo para a tomada de decisões deste órgão (ARAGÃO; ARAÚJO, 2015). Por sua vez, Borges, Libório e Haddad (2019), observaram que a ausência de informações espaciais torna mais onerosa a realização de análises investigativas, prolongando o tempo de investigação e dificultando a proposição de estratégias, considerando a atuação do Ministério Público do Rio Grande do Norte.

Ao longo dos anos, a instituição tem evoluído de modo a garantir maior atuação em prol do meio ambiente, contando com Promotorias Especializadas, Centros de Apoio Operacionais e Núcleos Ambientais. A exemplo, o Ministério Público do Estado de Mato Grosso do Sul (MPMS), desde 2007, conta com um Núcleo de Geotecnologias (NUGEO), vinculado ao Centro de Apoio Operacional das Promotorias de Justiça do Meio Ambiente, da Habitação e Urbanismo e do Patrimônio Histórico e Cultural (CAOMA). Dentre suas

atribuições, este núcleo busca oferecer auxílio as Promotorias de Justiça no cumprimento da legislação ambiental, com ênfase na questão dos desmatamentos e queimadas ilegais, bem como irregularidades em áreas de reserva legal e preservação permanente (MPMS, 2011).

No ano de 2016, o NUGEO lançou um programa para a detecção de desmatamento ilegal de vegetação nativa no Estado de Mato Grosso do Sul, denominado Programa DNA Ambiental, no qual as supressões nos biomas Cerrado, Mata Atlântica e Pantanal são identificadas através da análise multitemporal via imagem de satélite (MPMS, 2019; MPMS, 2020). Após a identificação, os dados são confrontados com as licenças de desmatamento emitidas pelos órgãos competentes, com intuito de evidenciar possíveis desmatamentos ilegais, resultando na emissão de pareceres aos órgãos ambientais e às Promotorias de Justiça para a tomada de providências e responsabilização dos infratores (GONÇALVES et al., 2021).

## 2.4 Aprendizado de máquina

Esta seção apresenta as definições dos principais conceitos e tecnologias utilizados neste trabalho relativos ao uso de aprendizagem de máquina.

### 2.4.1 Segmentação semântica

A atividade de atribuir uma classe para cada pixel de uma imagem, de forma que cada região da imagem esteja rotulada com uma classe diferente, é conhecida como segmentação semântica (WANG; LIU; XU, 2017). A Figura 3 apresenta uma representação desta atividade exibindo a imagem original ao lado esquerdo e sua versão segmentada ao lado direito.

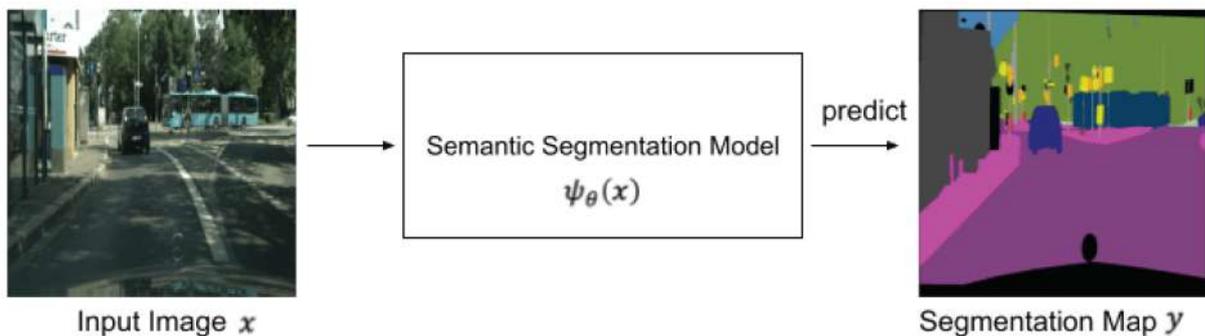


Figura 3 – Exemplo de segmentação (Figura reproduzida de (KIYA et al., 2022) Licença: CC BY 4.0 DEED).

Recentemente as *Fully Convolutional Neural Networks* (FCN) vêm se destacando para a realização desta tarefa (SHELHAMER; LONG; DARRELL, 2017; LONG; SHELHAMER; DARRELL, 2014). A princípio uma FCN é muito parecida com uma Rede

Neural Convolutacional, sendo que para a atividade de classificação é possível simplesmente substituir as camadas densamente conectadas ao final da rede por filtros convolucionais 1x1, dimensionando a quantidade de mapas de ativação, aumentando a capacidade da rede para resolver problemas não lineares e obtendo uma rede completa apenas com camadas convolucionais.

Além da atividade de classificação, as *Fully Convolutional Neural Networks* também podem ser utilizadas para segmentação semântica, considerando a possibilidade de aplicação de técnicas de *upsampling* ao final da rede para garantir que saída da última camada tenha a mesma resolução da imagem original, já com cada pixel rotulado com sua respectiva classe (XU; LIAN, 2018).

O grande desafio é garantir que durante a aplicação das camadas de *upsampling* a rede seja capaz de preservar todas as características espaciais da imagem original, além de desenvolver a capacidade de reconhecer com precisão a fronteira entre objetos que podem existir em diferentes escalas. Ao longo do tempo várias técnicas foram experimentadas na expectativa de melhorar a capacidade deste tipo de arquitetura, como o uso de *hierarchical networks* (TAO; SAPRA; CATANZARO, 2020), *encoders* e *decoders* (RONNEBERGER; FISCHER; BROX, 2015), *skip connections* (HE et al., 2015), *atrous convolutions* (CHEN et al., 2018), *pyramid pooling* (ZHAO et al., 2016) e mecanismo de atenção (CHEN et al., 2015).

## 2.4.2 Segmentação semântica para detecção de mudanças

Um dos principais desafios para aplicação de técnicas de segmentação semântica em séries temporais de imagens de satélite está no fato de a camada de entrada das principais arquiteturas ter sido projetada para o processamento de uma única imagem.

Considerando o problema de detecção de mudanças em séries temporais de imagens com treinamento supervisionado, Shi et al. (2020) consideram que as principais soluções podem ser classificadas em 3 abordagens diferentes:

### 1. *Single-Stream*

Uma abordagem para migração de uma solução de segmentação semântica para a tarefa de detecção de mudanças é a aplicação da técnica de Fusão Precoce, onde ambas imagens são concatenadas de modo a se ajustarem a arquitetura que espera uma única imagem de entrada. Um possível problema com esta abordagem é o fato de que as camadas iniciais da rede com Fusão Precoce podem perder informações representativas para a reconstrução da imagem original, podendo resultar em um mapa de mudanças com falhas na definição das fronteiras entre as classes (ZHANG et al., 2020).

## 2. *Double-Stream*

As técnicas para esta abordagem se baseiam em arquiteturas de Redes Neurais Convolucionais que recebem como entrada as duas imagens da série temporal separadamente. Para esta atividade temos duas abordagens principais: a Fusão Tardia e as redes Siamesas.

As redes Siamesas podem ser classificadas de acordo com sua estrutura de compartilhamento dos pesos. As redes puramente siamesas são aquelas que extraem as características comuns entre as duas imagens compartilhando os pesos entre elas, enquanto as pseudo-siamesas extraem as características de cada imagem separadamente, resultando em um grande aumento na quantidade de parâmetros treináveis (SHI et al., 2020).

Maretto et al. (2020) sugeriram uma metodologia de fusão para considerar o aspecto temporal das imagens. Neste caso, após a concatenação das imagens, seria executado um filtro de convolução 1x1 com a mesma quantidade de filtros que o número de canais original de cada imagem. A expectativa é que estes filtros aprendam a reconhecer as alterações entre as imagens. Esta técnica pode ser aplicada logo após a entrada da rede, e neste caso seria considerada uma Fusão Precoce ou após cada bloco de convolução, quando seria considerada uma Fusão Tardia.

## 3. *Multi-model integrated*

Este método possui uma estrutura híbrida, similar ao *Double-Stream*, mas utiliza diferentes tipos de arquiteturas possibilitando a divisão do processo de treinamento em vários estágios (SHI et al., 2020).

### 2.4.3 U-NET

Entre as redes neurais artificiais mais conhecidas está a U-Net, inicialmente proposta por Ronneberger, Fischer e Brox (2015) para segmentação na área da medicina, esta arquitetura obteve resultados expressivos nas mais diversas aplicações.

A arquitetura foi proposta utilizando a técnica de *autoencoder*, como é possível observar na Figura 4, ao lado esquerdo acontece a contração, também chamado de *encoder*, onde a rede aplica convoluções e faz o *downsampling* utilizando operações de *pooling*, enquanto o lado direito é responsável pela expansão, também chamado de *decoder*, utilizando convoluções e aplicando *upsampling* com operações de convoluções transpostas.

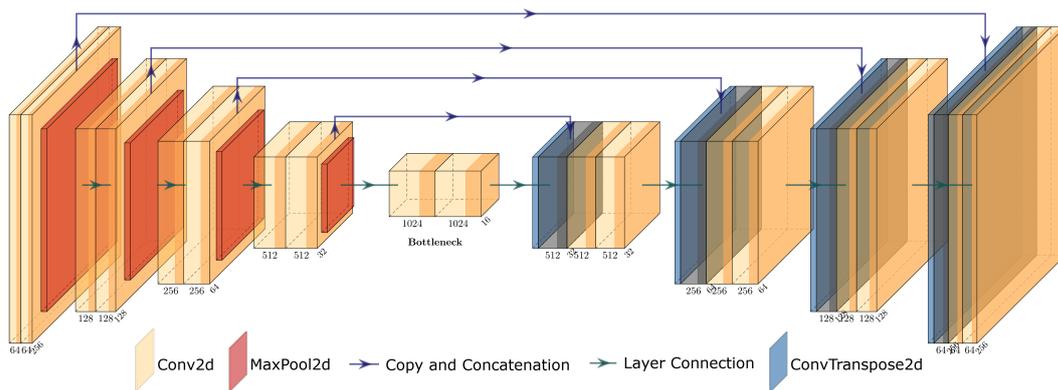


Figura 4 – Arquitetura da U-Net. (Figura produzida utilizando a ferramenta PlotNeuralNet (IQBAL, 2018))

A convolução é uma técnica já estabelecida na literatura mas a convolução transposta ainda não foi exhaustivamente explicada. Enquanto no processo de convolução um filtro, ou *kernel*, nas dimensões N por M é aplicado a um fragmento com as mesmas dimensões do vetor de entrada e produz um único resultado de saída, a convolução transposta aplica um filtro com dimensões N por M (Figura 5/B) em uma única posição do vetor de entrada (Figura 5/A) e produz uma saída ampliada (Figura 5/C) (NOH; HONG; HAN, 2015). Embora esta técnica tenha sido chamada de deconvolução, é preciso cautela na utilização deste termo (IM et al., 2016) considerando que ele pode, equivocadamente, levar a conclusão de que o processo de convolução tenha sido desfeito.

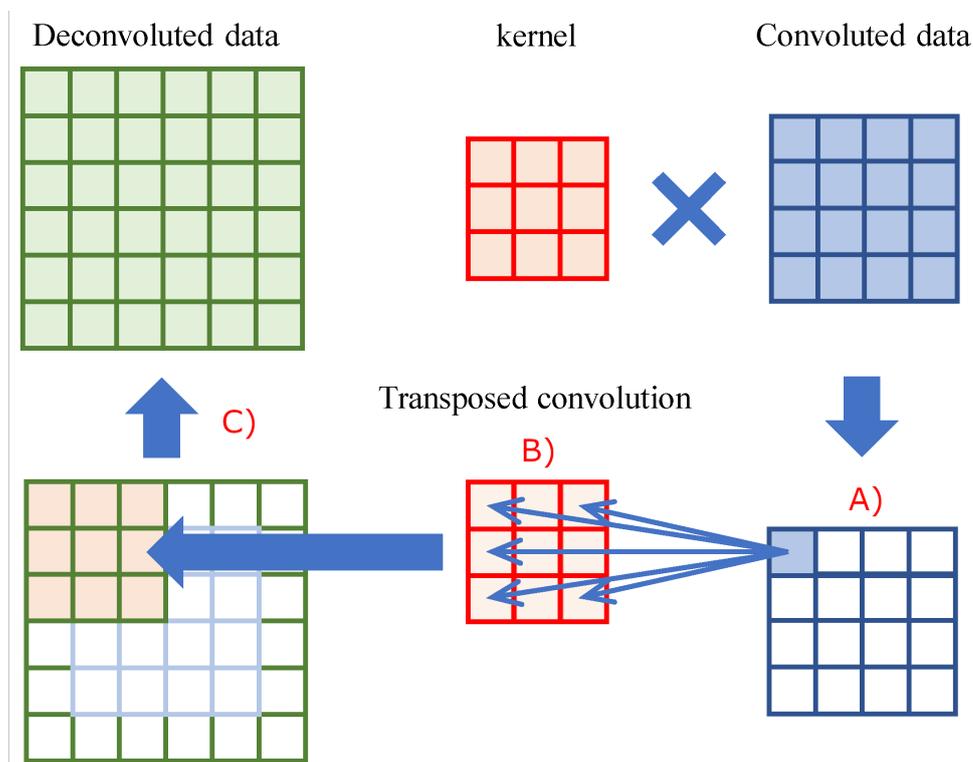


Figura 5 – Ilustração de uma convolução transposta (Figura adaptada de (MACHIDA; NAMBU; WADA, 2023) Licença: CC BY 4.0 DEED)

Para preservar características espaciais da imagem original a rede utiliza *skip connections*, concatenando o resultado das convoluções durante o processo de contração com o resultado da operação de convolução transposta da respectiva camada durante o processo de expansão.

#### 2.4.4 DeepLab-based Change Detection (DLCD)

A *DeepLab-based Change Detection* (DLCD), proposta por [Andrade et al. \(2020\)](#), é uma arquitetura para segmentação semântica baseada na DeepLabv3+ ([CHEN et al., 2018](#)) adaptada para a atividade de detecção de desmatamento.

Um grande desafio na utilização de FCNs para segmentação semântica está na imagem de saída da rede que deve ter as mesmas dimensões da imagem de entrada. Durante a execução da rede, as camadas de convoluções e *poolings* tornam as informações mais densas, que por um lado favorece a utilização de muitas camadas ocultas pois diminui a quantidade de parâmetros, mas por outro lado exige que a rede execute operações de *upsampling* para dimensionar a camada de saída, que pode ocasionar na perda de qualidade.

Outra questão importante é que objetos de uma mesma classe podem ser apresentados na imagem em escalas diferentes e a rede precisa aprender a reconhecê-los mesmo que apareçam em uma escala nunca utilizada durante o treinamento.

Para solucionar estas questões, [Chen et al. \(2018\)](#) propõem a utilização de *atrous convolution*, ou convolução dilatada, juntamente com o *Atrous Spatial Pyramid Pooling* (ASPP) e *Depthwise separable convolution* em uma estrutura de *Encoder-decoder*.

1. *Atrous convolution*, ou convolução dilatada, é uma forma de convolução que tem o objetivo de aumentar o campo de visão, ou *field-of-view*, de cada filtro sem a necessidade de utilização de filtros muito grandes e, conseqüentemente, sem o aumento na quantidade de parâmetros treináveis ou aumento de processamento.
2. *Atrous Spatial Pyramid Pooling* (ASPP) é uma estrutura parecida com a técnica de *pyramid pooling* proposta por [Zhao et al. \(2016\)](#) para a PSPNet, porém, no lugar da execução de convoluções paralelas em uma escala de *grids*, são utilizadas convoluções dilatadas paralelas com o parâmetro *rate* alterado.
3. *Depthwise separable convolution* - Enquanto convoluções de modo geral atuam em todas as camadas de entrada, uma *depthwise convolution* atua em cada camada, ou grupo de camadas, separadamente. Ao aplicar esta operação seguida de uma camada de convolução 1x1 é possível obter um resultado semelhante ao de uma convolução comum reduzindo significativamente a quantidade de parâmetros treináveis e o custo computacional.

### 2.4.5 Hierarchical Multi-Scale Attention for Semantic Segmentation

Esta arquitetura, proposta por Tao, Sapra e Catanzaro (2020), ocupa atualmente um dos primeiros lugares no ranking *Cityscapes test*, que é um popular *dataset* utilizado para comparação de resultados de segmentação semântica contendo cerca de 5 mil exemplos de imagens anotadas perfeitamente e 20 mil exemplos de imagens anotadas de maneira irregular, distribuídas entre 30 classes diferentes.

A técnica foi baseada na proposta de *Multi-Scale Attention* (CHEN et al., 2015), que tem em sua arquitetura duas entradas de uma mesma imagem em diferentes escalas e utiliza o mecanismo de atenção para determinar qual região de cada imagem deve receber mais responsabilidade relativa a saída esperada, com a estrutura adaptada para atuar de forma hierárquica. Na Figura 6 é possível perceber que durante a utilização deste modelo a imagem em escala menor fica responsável por determinar a segmentação dos grandes objetos (círculo pontilhado rosa) enquanto a imagem com escala maior pode focar na segmentação dos objetos menores (círculo pontilhado verde).

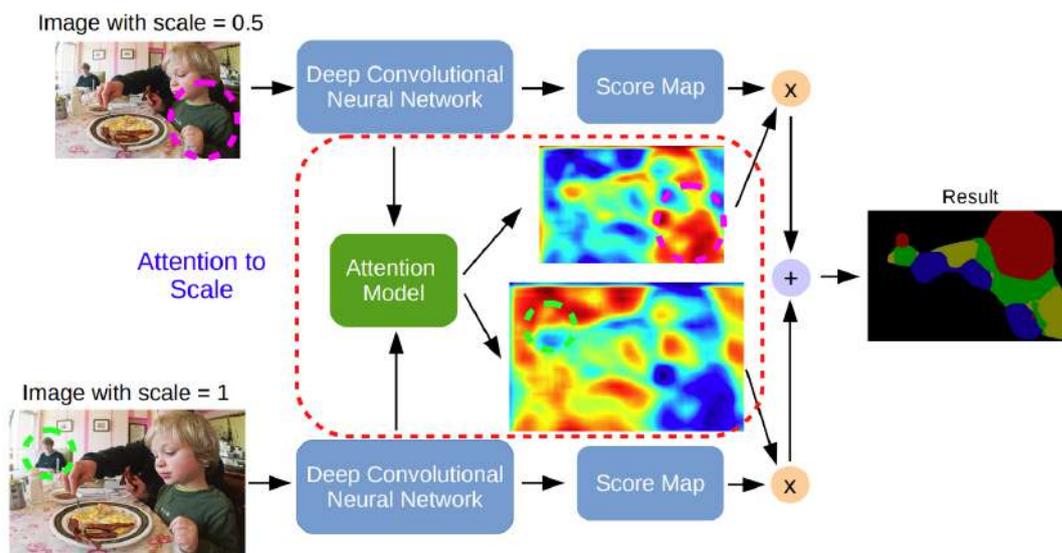


Figura 6 – Ilustração do modelo *Multi-Scale Attention* (Figura reproduzida de (CHEN et al., 2016) IEEE copyright line © 2016).

### 2.4.6 Métricas de avaliação

Métricas de avaliação são técnicas para estabelecer o desempenho relativo de diferentes modelos a partir de uma determinada perspectiva, possibilitando a comparação entre os resultados. Dependendo da natureza do problema abordado diferentes métricas podem ser utilizadas, nesta seção teremos uma breve descrição de três diferentes métricas que podem ser utilizadas para a tarefa de segmentação semântica.

*Intersection over Union* (IoU) é uma métrica de avaliação frequentemente utilizada em casos de segmentação semântica, obtida pela divisão da área de intersecção entre

o objeto predito e o objeto esperado, que é basicamente a parte que a rede rotulou corretamente, dividido pela soma da área total do objeto predito com a área total do objeto esperado. A média desta métrica é comumente utilizada pelos *benchmarks* de detecção de mudanças para comparação entre os melhores resultados.

*Pixel accuracy* é uma métrica muito simples que basicamente divide a quantidade de *pixels* que o modelo acertou pela quantidade total de *pixels* da imagem. É necessário ter cuidado com a utilização dessa métrica caso o conjunto de dados seja muito desbalanceado, afinal o resultado apresentado pode parecer positivo mesmo que a rede esteja classificando todos os *pixels* como sendo da classe com maior ocorrência, ou seja, errando completamente a previsão para as outras classes.

*F1-score* é uma métrica de avaliação que utiliza uma média harmônica entre os valores de *Precision* e *Recall*. A medida de avaliação *Precision*, relaciona o número de exemplos verdadeiros positivos (VP) com o total de exemplos avaliados como positivos, ou seja, a soma dos falsos positivos (FP) e verdadeiros positivos (VP), conforme demonstra a equação a seguir.

$$\text{Precision} = \frac{VP}{FP+VP}$$

Já a medida de avaliação *Recall* representa a qualidade da classificação da classe positiva ao relacionar o número de exemplos verdadeiros positivos (VP) com o seu somatório ao número de exemplos falsos negativos (FN), de acordo com a equação a seguir.

$$\text{Recall} = \frac{VP}{FN+VP}$$

Desta forma, o cálculo da F1 score pode ser definido conforme qualquer das seguintes equações:

$$\text{F1 Score} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

$$\text{F1 Score} = \frac{2 * VP}{2 * VP + FP + FN}$$



## 3 Trabalhos Relacionados

Diversos estudos recentes indicam resultados promissores em relação à eficácia das redes neurais profundas para a atividade de detecção de mudanças (LUO et al., 2021; SONG; JIANG, 2021; CHEN; SHI, 2020; GUO et al., 2018; DAUDT; SAUX; BOULCH, 2018). Especificamente para a tarefa de detecção de desmatamento pelo mundo, tivemos uma quantidade considerável de trabalhos (SOLÓRZANO et al., 2023; ZHAO et al., 2022; ISAIENKOV et al., 2021; LEE et al., 2020; KATARKI et al., 2019; KHAN et al., 2017). Neste capítulo apresenta-se a descrição das recentes utilizações de redes neurais profundas para a detecção de desmatamento nos biomas brasileiros, com especial foco para os trabalhos relacionados aos biomas Amazônia, Cerrado e Mata Atlântica.

### 3.1 Amazônia

Nos últimos anos, algumas pesquisas apontaram vantagens no uso de redes neurais profundas para detecção de desmatamento no bioma Amazônia (Tabela 1). Mesmo quando utilizados com diferentes metodologias, conjunto de dados e aplicações, ainda é possível observar que algumas arquiteturas de redes neurais profundas seguem uma tendência evidente.

Ortega et al. (2019) estão entre os primeiros a estudarem técnicas de aprendizado profundo para detecção de desmatamento na Amazônia, comparando arquiteturas *Early Fusion* (EF) e *Siamese Convolutional Neural Network* (S-CNN) com a tradicional *Support Vector Machine* (SVM) para classificação de fragmentos de imagens. O conjunto de dados utilizado foi um par de cenas do satélite Landsat-8, ambos cortados para as dimensões  $1100 \times 2600$  pixels cada enquanto adicionaram uma camada de *Normalized Difference Vegetation Index* (NDVI) às 7 camadas originais do satélite, ficando com um total de 8 camadas espectrais por imagem. Este *dataset* foi dividido em 15 partes e então separado em 3 grupos para treinamento, validação e teste. O grupo de treinamento contou com o aumento de dados para balancear a quantidade de exemplos entre as classes. As imagens foram fragmentadas em trechos de  $15 \times 15$  pixels para utilização na camada de entrada das redes. A SVM ficou com o pior resultado, destacando a diferença entre as tecnologias tradicionais e as Redes Neurais Profundas para este tipo de aplicação. Entre as Redes Neurais Profundas o resultado foi muito próximo, as quais conseguiram 0.65 de F1-Score utilizando a S-CNN, apenas 0.03 melhor em comparação com o resultado obtido pela EF.

Utilizando um conjunto de dados e metodologia similares, Andrade et al. (2020) também testaram as arquiteturas EF e S-CNN para classificação dos fragmentos de imagem enquanto introduziram uma variação da DeepLabv3+(CHEN et al., 2018) para seg-

Tabela 1 – Comparação entre resultados recentes para detecção de desmatamento na Amazônia utilizando aprendizado profundo.

Autores	Ano	Satélite	Arquitetura	Loss	F1	
Ortega et al.	2019	Landsat-8	SVM		0.5200	
			EF CNN		0.6200	
			<b>S-CNN</b>		<b>0.6500</b>	
Andrade et al.	2020	Landsat-8	EF CNN	WFL	0.6320	
			S-CNN	WFL	0.6290	
			<b>DLCD (DeepLabv3+)</b>	WFL	<b>0.7180</b>	
Bem et al.	2020	Landsat-8	Random Forest		0.8014	
			MLP		0.8926	
			<b>Resunet</b>		<b>0.9432</b>	
			U-Net		0.9112	
			SharpMask		0.9223	
Maretto et al.	2020	Landsat-8	U-Net	WCE	0.9448	
				ASD	0.9417	
			EF U-Net	WCE	0.9427	
				ASD	0.9388	
			<b>LF U-Net</b>	WCE	<b>0.9471</b>	
	ASD	0.9460				
Torres et al.	2021	Landsat-8	SegNet	WCE	0.596	
			U-Net	WCE	0.674	
			<b>ResU-Net</b>	WCE	<b>0.707</b>	
			FC-DenseNet	WCE	0.691	
			DeepLabv3+ (Xception)	WCE	0.640	
	Sentinel 2	2021	Sentinel 2	DeepLabv3+ (MobileNetV2)	WCE	0.604
				SegNet	WCE	0.633
				U-Net	WCE	0.682
				ResU-Net	WCE	0.702
				<b>FC-DenseNet</b>	WCE	<b>0.707</b>
DeepLabv3+ (Xception)	WCE	0.651				
DeepLabv3+ (MobileNetV2)	WCE	0.615				

mentação semântica, neste caso com algumas pequenas alterações na metodologia. Para o trabalho de segmentação as imagens foram fragmentadas em trechos de 64 x 64 pixels e o balanceamento do conjunto de dados de treinamento não foi resolvido com o aumento de dados mas utilizando uma *weighted focal loss* (WFL). Para as técnicas de classificação, os resultados obtidos corroboram com as conclusões apresentadas pelo estudo anterior, com F1-Score próximo a 0.65 em ambas arquiteturas. Neste mesmo trabalho a solução proposta para segmentação semântica se sobressaiu apresentando o resultado de 0.71 de F1-Score.

Bem et al. (2020) apresentaram uma metodologia distinta e tiveram como objetivo a comparação de resultados para segmentação semântica entre cinco arquiteturas: *random forest* (RF); *multilayer perceptron* (MLP); Resunet; Unet; e Sharpmask. O conjunto de dados foi construído a partir de imagens de três diferentes regiões da Amazônia registradas pelo satélite Landsat-8, sendo que duas destas imagens foram utilizadas para treinamento e validação enquanto a última imagem foi utilizada para testes. Para aplicação como entrada das redes neurais profundas as imagens foram fragmentadas em trechos contendo 200 x 200 *pixels*. Estes resultados também apontaram para vantagem na utilização de redes neurais profundas em relação as técnicas tradicionais de aprendizado de máquina, mostrando que a arquitetura Resunet se destacou com 0.94 de F1-Score, enquanto a RF apresentou admiráveis 0.80 com a mesma métrica, indicando uma provável diferença entre o conjunto de dados deste estudo e os conjuntos de dados utilizados pelas pesquisas supracitadas.

Maretto et al. (2020) apontaram como objetivo não somente analisar a efetividade dos algoritmos em um ambiente controlado mas a automatização de todo o processo de classificação. Neste trabalho uma U-Net foi utilizada como base enquanto comparada com uma EF e uma *Late Fusion (LF) Spatio-Temporal U-Net*, cada uma delas foi utilizada com as funções de perda *Average Soft Dice* (ASD) e *Weighted Cross-Entropy* (WCE). Em seu melhor resultado o F1-Score de 0.94 foi alcançado utilizando a arquitetura U-Net com *Late Fusion* em conjunto com a função de perda WCE, mas deve-se observar que todas as arquiteturas selecionadas alcançaram resultados extraordinariamente parecidos, o que pode indicar que o conjunto de dados utilizado apresentava baixa complexidade ou pouco ruído.

Por sua vez, Torres et al. (2021) explorou a utilização das redes convolucionais U-Net, ResU-Net, SegNet, FC-DenseNet e duas variantes do DeepLabv3+, com imagens dos satélites Landsat-8 e Sentinel 2, associada a aplicação de auditoria dos resultados. Os melhores resultados foram obtidos com as redes ResU-Net e FC-DenseNet, com os valores da métrica F1-Score iguais a 0.70 para ambos os casos.

Pesquisas mais recentes, como as desenvolvidas por Park et al. (2023) e Alshehri, Ouadou e Scott (2024), revelam que o emprego de redes neurais profundas para a detecção

de desmatamento no bioma Amazônia continua a trilhar caminhos promissores, com a utilização de arquiteturas baseadas em *Transformer*. O primeiro alcançou o resultado de 0.88 em F1-Score, com a rede Mask2Former. Já o segundo obteve o melhor resultado de 0.90 em F1-Score, com a rede ChangeFormer.

## 3.2 Cerrado e Mata Atlântica

Para os biomas Cerrado e Mata Atlântica a quantidade de pesquisas relacionadas reduz consideravelmente em comparação com o bioma Amazônia. Em uma pesquisa utilizando a frase “deforestation deep learning amazon” na ferramenta *Google Scholar* é possível obter aproximadamente 30 mil resultados, enquanto a mesma pesquisa substituindo a palavra “Amazon” por “Cerrado”, exibe uma quantidade de resultados expressivamente inferior, com cerca de 4 mil resultados.

Um grande desafio deste tipo de trabalho é a combinação da implementação de arquiteturas para segmentação semântica com o uso de séries temporais. [Parente et al. \(2019\)](#) utilizaram uma U-net e uma LSTM separadamente e, embora tenham conseguido resultados interessantes utilizando uma técnica de múltiplos modelos integrados, o uso de dados privados do satélite PlanetScope, que possui resolução muito superior ao Landsat-8, dificulta a comparação com o resultado dos demais artigos.

[Adarme et al. \(2020\)](#) publicaram um estudo analisando a utilização de técnicas de *Deep Learning* para detecção de desmatamento nos biomas Amazônia e Cerrado aplicadas para a atividade de classificação em imagens do satélite Landsat-8 e obtiveram resultados similares para ambos os biomas, chegando a 0.77 de F1-Score com uma *Early Fusion Convolutional Network*. Ainda que esta pesquisa tenha avaliado diversos modelos de *Deep Learning* utilizando dados referentes ao bioma Cerrado, restou dúvida quanto ao desempenho relativo a este bioma quando analisado em uma solução implementada para a atividade de segmentação semântica, e também quanto ao desempenho das demais arquiteturas que obtiveram resultados promissores em outros biomas, como a DeepLabv3+ proposta por [Andrade et al. \(2020\)](#), e a LF U-Net proposta por [Maretto et al. \(2020\)](#).

Em um estudo mais recente analisando técnicas de Deep Learning para o bioma Cerrado, publicado por [Matosak et al. \(2022\)](#), foram feitos experimentos utilizando dados dos satélites Landsat-8 e Sentinel 2, assim como imagens de diferentes estados brasileiros: Bahia e Mato Grosso. Para o estado de Mato Grosso, sendo o mais próximo do atual estudo, foi obtido o F1-Score de 0.5428 na melhor de suas abordagens para as imagens do satélite Sentinel 2, enquanto para o satélite Landsat-8 o melhor F1-Score obtido considerando as imagens do mesmo estado foi de 0.7128.

A tabela 2 compara as abordagens definidas por três trabalhos relacionados ([MATOSAK et al., 2022](#); [ADARME et al., 2020](#); [PARENTE et al., 2019](#)) que utilizaram

Tabela 2 – Comparação das abordagens utilizadas para detecção de desmatamento no bioma Cerrado.

<b>Autor</b>	<b>Satélite/Região</b>	<b>Entrada</b>	<b>Loss Function</b>	<b>Modelos</b>
<a href="#">Parente et al. (2019)</a>	Planet Scope / Goiás	RGB + NIR	BCE	Random Forest, LSTM, U-Net
<a href="#">Adarme et al. (2020)</a>	Landsat 8 / Maranhão	RGB + NIR + Coastal / Aerosol + SWIR-1 + SWIR-2 + NDVI	BCE	EF, SN, CSVM
<a href="#">Matosak et al. (2022)</a>	Landsat 8, Sentinel 2 / Bahia, Mato Grosso	RGB + NIR + SWIR-1 + SWIR-2 + NDVI + EVI	CE	LSTM, U-Net

técnicas de Aprendizado Profundo para detecção de desmatamento no bioma Cerrado.

Enquanto estes estudos foram importantes para evidenciar a eficácia no uso de redes neurais profundas para a atividade de detecção de desmatamento, ainda há espaço para novos trabalhos relacionados considerando não somente a pequena quantidade de estudos relacionados aos biomas Cerrado e Mata Atlântica, mas também a necessidade de deslindar questões importantes que ainda restaram abertas, como a influência da qualidade do conjunto de dados e da falta de conjuntos de dados abertos para realização de pesquisas e experimentações sobre o tema.



## 4 Metodologia

A proposta deste trabalho foi automatizar o processo de classificação de imagens de satélite realizado pela equipe do NUGEO, quanto a ocorrência de desmatamentos, disponibilizando um modelo de inteligência artificial treinado, com a capacidade mapear áreas desmatadas em qualquer cena dos biomas Cerrado e Mata Atlântica, utilizando o melhor resultado obtido nesta pesquisa.

Insta salientar que entre as técnicas recentemente utilizadas para resolução de problemas de detecção de desmatamento em séries temporais de imagens de satélite, destacam-se a classificação e a segmentação semântica, as quais têm apresentado resultados em F1 score e acurácia semelhantes, porém, suas saídas distinguem-se. Desta forma, considerando que a precisão na detecção de fronteiras das técnicas aplicadas é um aspecto relevante aos trabalhos desenvolvidos pela equipe do MPMS, nesta pesquisa foram investigadas técnicas de segmentação semântica, considerando seus melhores resultados neste quesito.

Para aplicação dessas técnicas, foi necessário recortar as cenas dos satélites em pequenas imagens, uma vez que o processamento de uma entrada com dimensões tão significativas, como a de uma cena em sua integralidade, exigiria elevada capacidade computacional para o processamento, dificilmente disponível. Desta forma, foram executados experimentos com o intuito de produzir um conjunto de dados apropriado para esta finalidade aplicando-se, em um primeiro momento uma metodologia para recorte das cenas disponibilizadas do satélite Landsat-8, obtendo-se o primeiro *dataset*. Posteriormente, foram executados novos experimentos com adaptações nessa metodologia, com objetivo de produzir uma segunda opção de *dataset*, com download e processamento de imagens do satélite Sentinel 2, conforme documentado nas seções a seguir.

### 4.1 Primeiro conjunto de dados – Landsat 8

Inicialmente foram disponibilizadas pela equipe do NUGEO duas imagens capturadas pelo satélite Landsat-8, referentes a agosto de 2015 e junho de 2018, correspondentes à uma área de 185 x 180 km cada. Respectivas cenas se sobrepõem aos seguintes municípios do Estado de Mato Grosso do Sul: 1-Anaurilândia, 2-Angélica, 3-Bataguassu, 4-Batayporã, 5-Campo Grande, 6-Deodápolis, 7-Douradina, 8-Dourados, 9-Fátima do Sul, 10-Glória de Dourados, 11-Itaporã, 12-Ivinhema, 13-Nova Alvorada do Sul, 14-Nova Andradina, 15-Novo Horizonte do Sul, 16-Ribas do Rio Pardo, 17-Rio Brillhante, 18-Santa Rita do Pardo, 19-Sidrolândia, 20-Taquarussu e 21-Vicentina, compreendidos entre os biomas Cerrado e Mata Atlântica, conforme é possível visualizar na Figura 7.

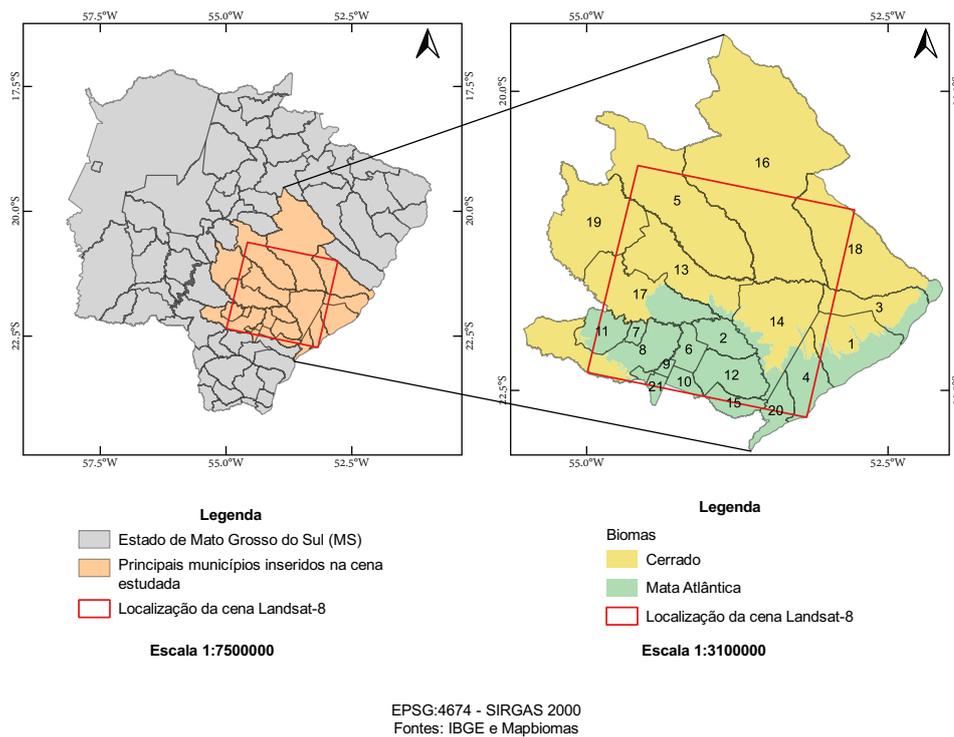


Figura 7 – Limite dos municípios no Estado de Mato Grosso do Sul.

Em conjunto com as cenas foi disponibilizado um arquivo em formato *Shapefile* contendo polígonos de todas as áreas que foram consideradas desmatadas no período comparado. Um exemplo destes polígonos de áreas desmatadas foi renderizado utilizando o programa QGIS e pode ser visualizado na Figura 8.

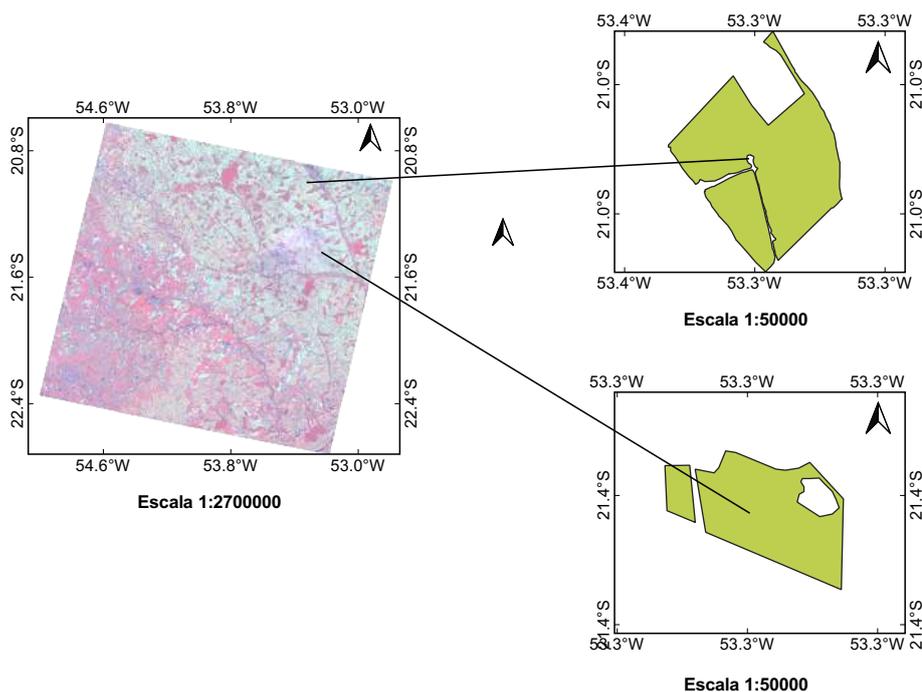


Figura 8 – Exemplo de polígonos de desmatamento disponibilizados no arquivo Shapefile.

A primeira cena, de agosto de 2015, apontava áreas sem cobertura vegetal, possivelmente desmatadas em períodos anteriores e que, portanto, não foram consideradas para este estudo. A outra cena, referente a junho de 2018, é responsável por apresentar áreas nas quais ocorreu a supressão da vegetação nativa entre o período de captura das duas imagens (2015 a 2018). Um exemplo desta situação pode ser observado na Figura 9.

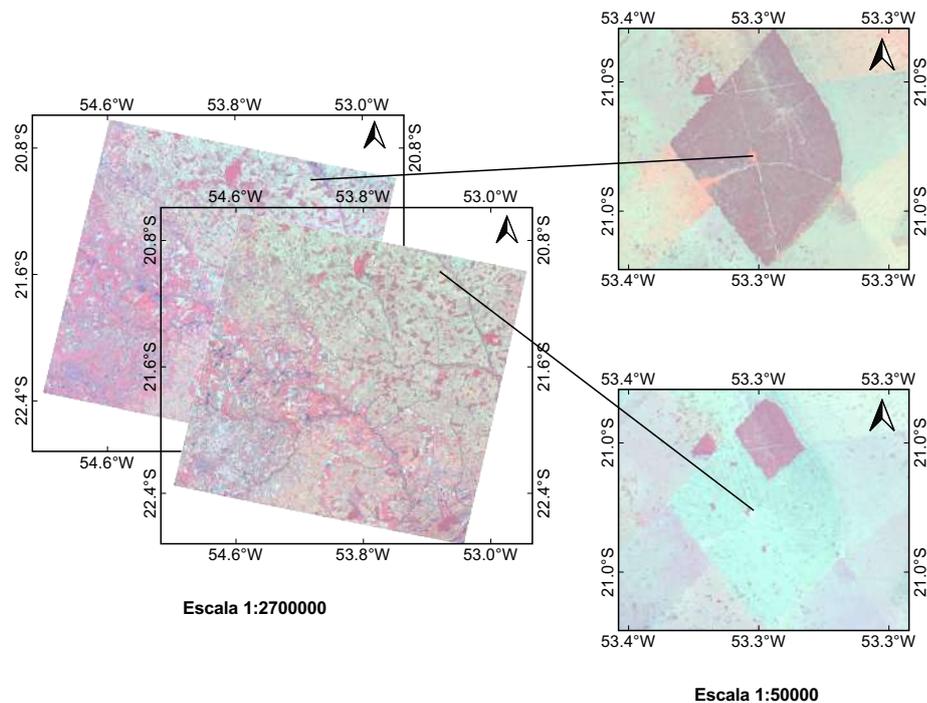


Figura 9 – Exemplo de área com desmatamento e sem desmatamento.

Ressalta-se que, além dos 7 espectros originalmente existentes nas imagens do satélite Landsat-8, foi gerado um novo espectro utilizando uma técnica denominada NDVI, consoante abordagem adotada por [Ortega et al. \(2019\)](#), resultando em imagens de entrada com 8 camadas.

Para o processamento dessas imagens, a primeira etapa realizada foi o recorte manual das duas cenas (Figura 10), cujas dimensões de entrada eram iguais a 15161 x 15401 x 8 (Largura x Altura x N<sup>o</sup> de camadas). Essas foram cortadas em pequenas imagens com as dimensões de 256 x 256 x 8 (Largura x Altura x N<sup>o</sup> de camadas), com intuito de otimizar a capacidade computacional utilizada em seu processamento, bem como selecionar as áreas com desmatamentos identificadas conforme arquivo *Shapefile* disponibilizado, descartando-se eventuais áreas não mapeadas que gerariam ruídos ao *dataset*, resultando em 340 recortes para cada cena, a serem divididos em conjuntos de treino, testes e validação.

Posteriormente, tendo em vista que a proporção dos *pixels* anotados como não desmatados foi consideravelmente maior que aqueles definidos como desmatados, foi necessário utilizar uma técnica para estratificar a divisão do *dataset* mantendo uma propor-

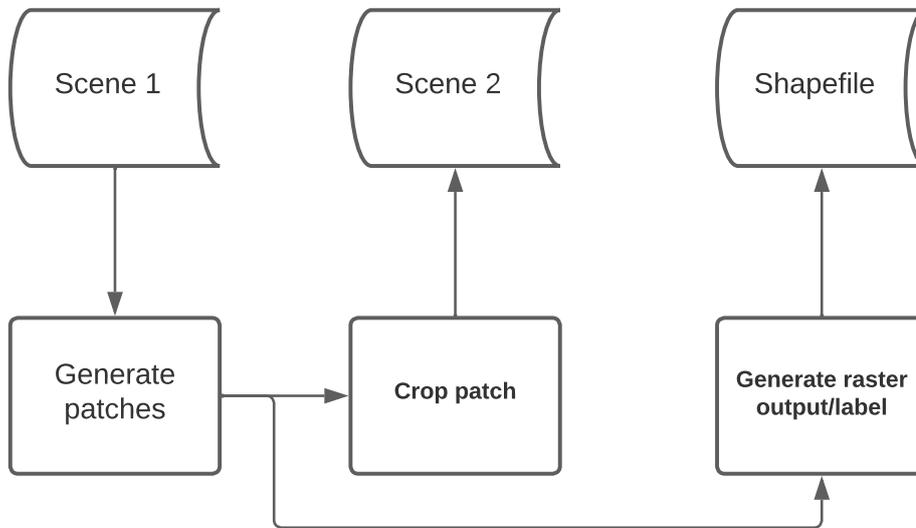


Figura 10 – Metodologia inicial - Processamento das cenas do satélite Landsat-8.

ção de áreas com maior desmate em todos os três conjuntos. Desta forma, o *dataset* foi dividido de modo aleatório e estratificado, adotando-se 70% dos recortes (238 recortes) para treinamento, 15% (51 recortes) para testes e 15% (51 recortes) para validação.

Para cada recorte das cenas utilizadas nesta pesquisa, foi gerada uma nova imagem *raster*, definindo cada um de seus *pixels* como desmatado ou não desmatado, com base em sua geolocalização e classe extraída do arquivo em formato *Shapefile* fornecido pelo NUGEO. Respectiva imagem foi utilizada como modelo de saída para a rede.

É importante salientar que, embora a etapa de recorte das cenas tenha sido realizada para viabilizar o seu processamento pelo modelo, esta tarefa também se demonstrou onerosa, demandando muito tempo e recursos, que inviabilizaram a realização de experimentos com diferentes dimensões de recorte neste trabalho.

Por fim, a última etapa do pré-processamento consistiu na normalização dos dados de cada camada, deixando seus valores em um intervalo entre 0 e 1.

Ademais, é relevante salientar que durante os experimentos preliminares de treinamento de algumas arquiteturas de redes neurais artificiais com este conjunto de dados, foi observado que esta metodologia resultou em muito ruído no *dataset*, que poderia ser mitigado por uma curadoria a ser realizada por especialistas para remoção dos exemplos defeituosos, enquanto provocaria a necessidade de obtenção de mais exemplos para treinamento.

## 4.2 Segundo conjunto de dados – Sentinel 2

Para aprimoramento do conjunto de dados criado anteriormente, foi disponibilizado pelo NUGEO outro arquivo *Shapefile* com novos polígonos de desmatamento em todo

o Estado de Mato Grosso do Sul, concentrados, em sua maioria, em áreas sobrepostas ao bioma Cerrado (Figura 11). Na ocasião foi solicitado pelo núcleo a alteração do escopo da pesquisa para que fossem consideradas imagens do satélite Sentinel 2, em decorrência de sua melhor resolução espacial e menor tempo de revisita.

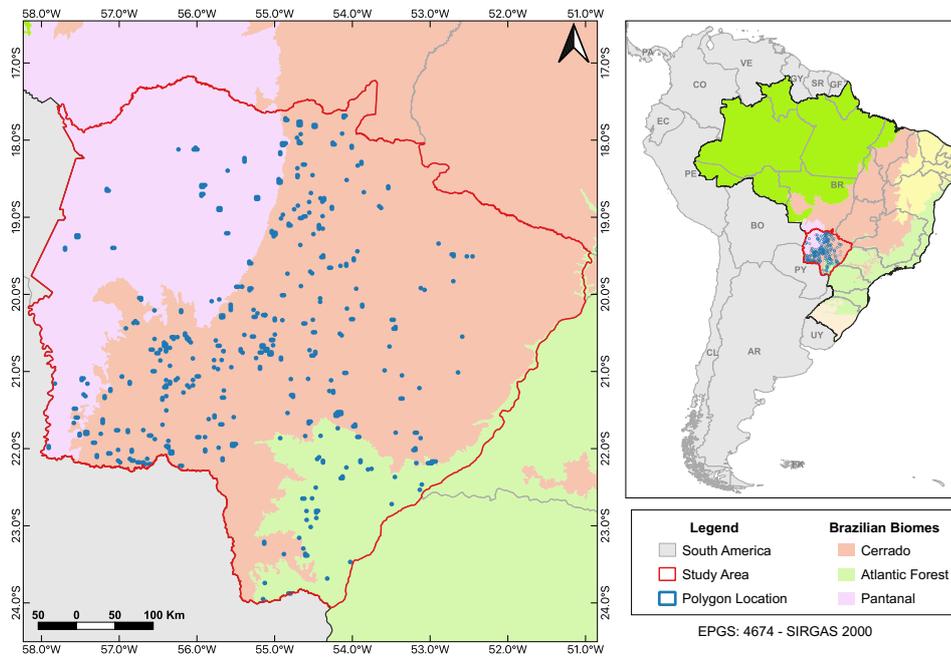


Figura 11 – Área de estudo relativa ao segundo conjunto de dados.

Diferentemente da natureza do primeiro conjunto de dados, no qual foram considerados os desmatamentos detectados em apenas duas cenas, durante o período entre duas datas definidas (agosto de 2015 e junho de 2018), os novos artefatos fornecidos possuíam polígonos de desmatamento detectados em diversas cenas e períodos distintos.

Nesse cenário, considerando o alto custo de processamento manual das cenas de imagens de satélites, tal como observado na geração do primeiro *dataset* com Landsat-8, foi desenvolvido um *framework* na linguagem python para o processamento automatizado dos dados provenientes das imagens do satélite Sentinel 2, utilizando APIs e infraestrutura em nuvem do projeto Planetary Computer da Microsoft ([Microsoft Open Source et al., 2022](#)).

Este *framework* realizou o download das cenas anteriores e posteriores a detecção dos desmatamentos, considerando as informações dos polígonos de desmatamento do arquivo *Shapefile* fornecido, buscando-se cenas com a menor porcentagem de nuvens, adotando-se uma janela de busca de 30 dias para cada cena. Em seguida, o *framework* efetuou o corte de cada cena baixada, nas dimensões 256 x 256 x 4 (Largura x Altura x N° de camadas), mantendo-se apenas os recortes com indicações de áreas desmatadas.

Ainda, foi gerada uma nova camada utilizando uma técnica denominada NDVI,

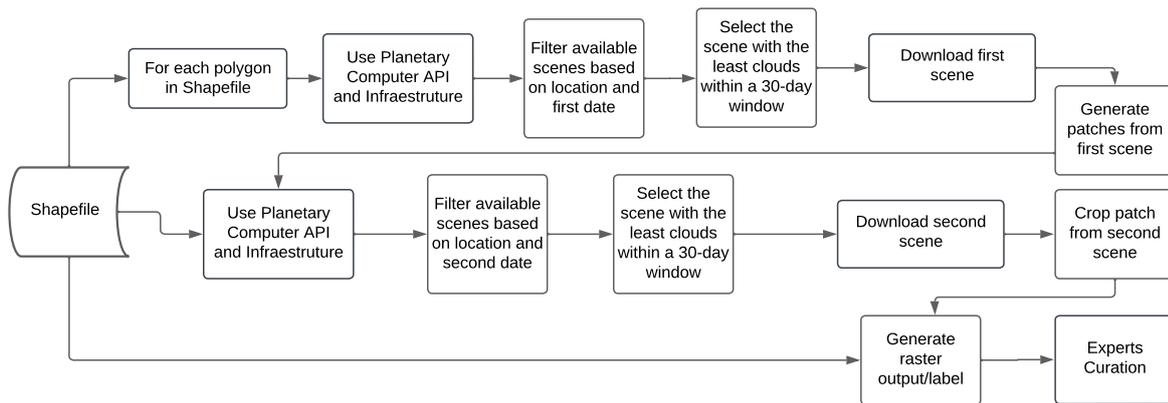


Figura 12 – Nova metodologia de preparação do dataset - Sentinel 2.

consoante abordagem adotada por [Ortega et al. \(2019\)](#), resultando em pares de imagens de entrada com 5 camadas (R+G+B+NIR+NDVI) para cada imagem.

Esta nova metodologia (Figura 12) proporcionou um conjunto de dados com maior quantidade de recortes, o que possibilitou a realização de uma curadoria especializada para redução dos ruídos, diferentemente do data *dataset* anterior.

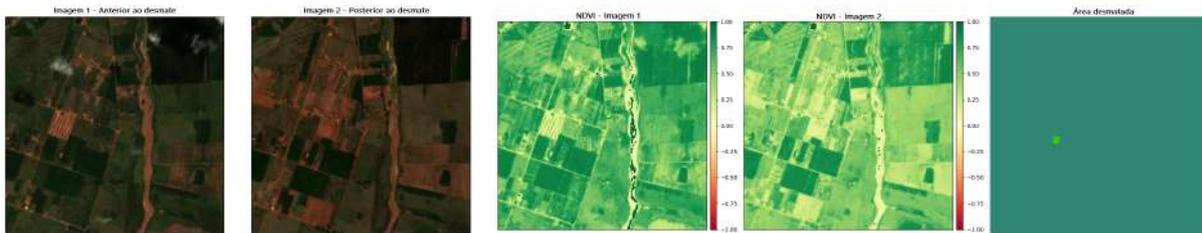


Figura 13 – Exemplo de recorte removido da amostra por conter possível desmate no canto superior direito não identificado.

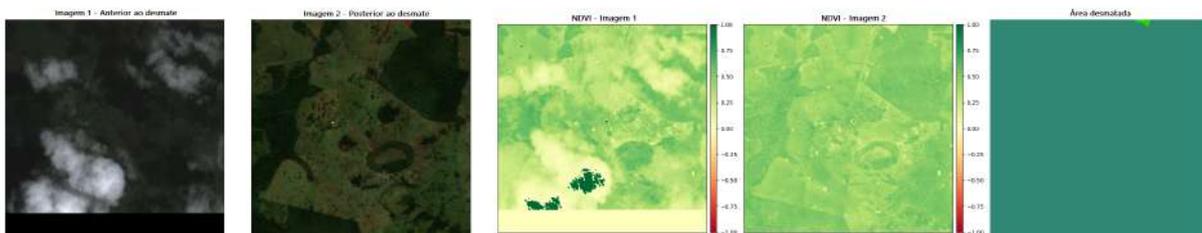


Figura 14 – Exemplo de recorte removido da amostra por conter falha na geração do recorte.

Durante o processo de curadoria, especialistas analisaram individualmente cada exemplo da amostra, destacando os exemplos considerados inadequados, por motivos diversos como, por exemplo, a ocorrência de possível área desmatada não identificada (Figura 13), quantidade excessiva de nuvens (Figura 15) e falha na geração do recorte (Figura 14). Esses casos foram removidos do conjunto de dados, mantendo-se apenas as imagens consideradas ideais (Figura 16).

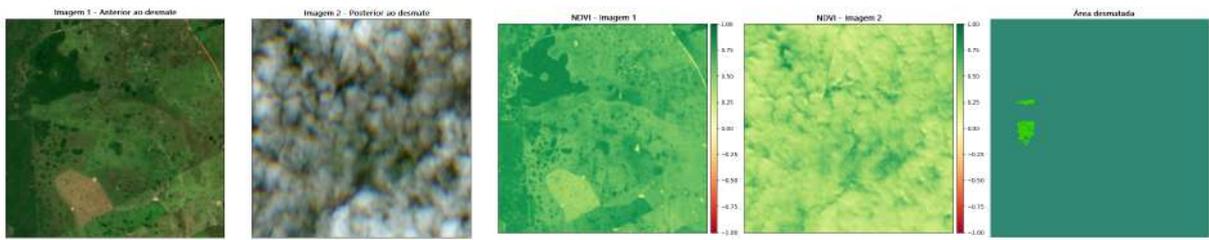


Figura 15 – Exemplo de recorte removido da amostra por conter quantidade excessiva de nuvens.

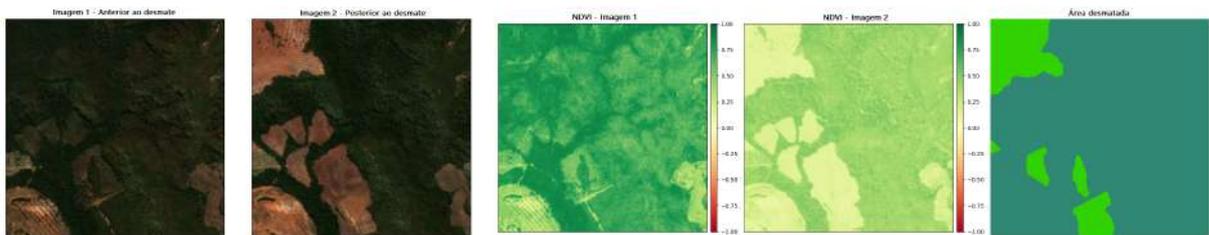


Figura 16 – Exemplo de recorte mantido na amostra por ter sido considerado ideal.

O novo conjunto de dados resultou em 514 recortes nas dimensões 256 x 256 *pixels*, os quais foram divididos em três amostras, conforme a metodologia de separação do dataset e treinamento dos modelos apresentada na Figura 17. A primeira, contendo 359 exemplos, definida como conjunto de treinamento e representando cerca de 70% da amostra completa. A segunda, com 78 exemplos, definida como conjunto de testes e, a última, contendo 77 exemplos, definida como conjunto de validação, ambas representando cerca de 15% do total de recortes.

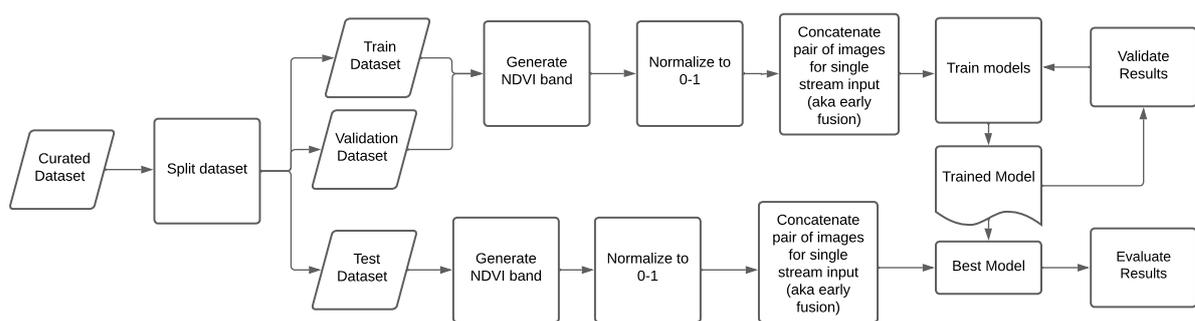


Figura 17 – Metodologia de treinamento.

Uma vez que a quantidade de pixels anotados para a classe positiva (com desmate) representa apenas 4.49% do total de pixels do conjunto, a divisão dos exemplos foi executada de maneira aleatória e estratificada, levando em conta o percentual de cobertura de área desmatada em cada imagem (Figura 18), com intuito de garantir uma distribuição uniforme das porções de área desmatada para todos os conjuntos (Figura 19).

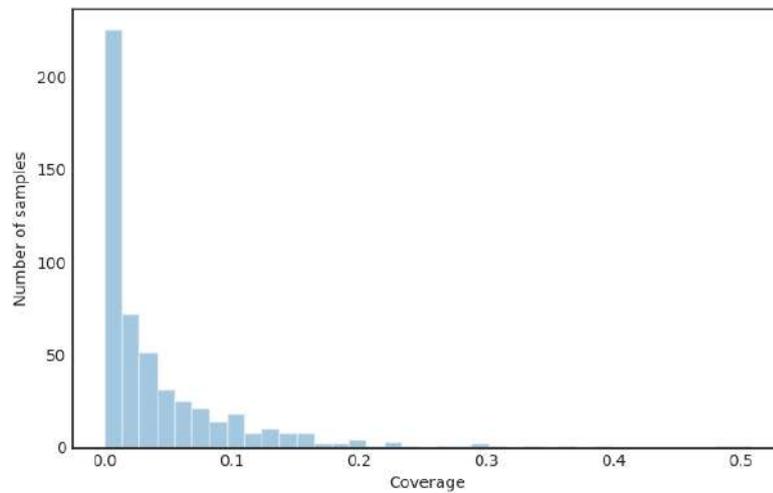


Figura 18 – Histograma - Quantidade de exemplos por percentual de cobertura de área desmatada.

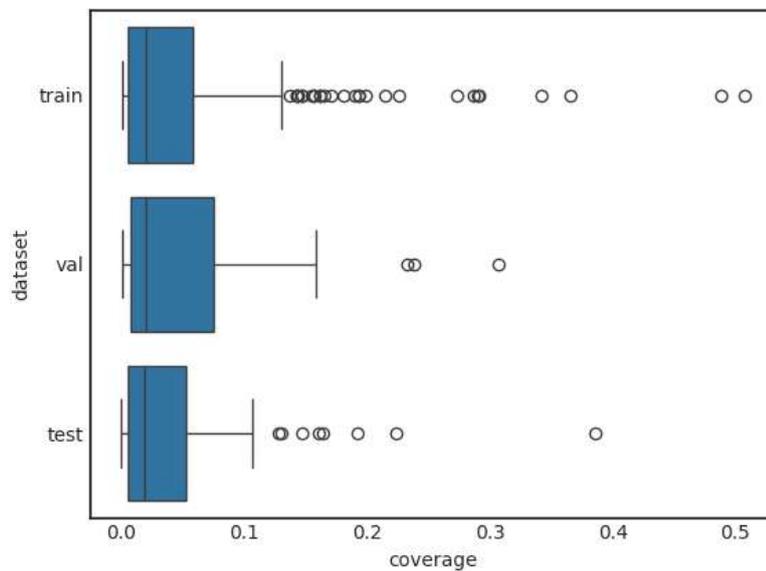


Figura 19 – Boxplot - Distribuição uniforme das porções de área desmatada para todos os conjuntos.

Na Figura 20 apresenta-se a distribuição dos conjuntos de dados, considerando as amostras de teste, validação e treinamento, nos limites da área de estudo. Destaca-se que estes estão localizados, em sua maioria, no centro do Estado de Mato Grosso do Sul, em área caracterizada como bioma Cerrado (IBGE, 2019).

### 4.3 Experimentos

Após a geração dos conjuntos de dados, foram executados testes e análises para avaliar os aspectos positivos e negativos de cada metodologia empregada para criação dos *datasets*, bem como foi analisada a qualidade dos resultados obtidos, possibilitando a definição das melhores práticas e a escolha do conjunto de dados mais adequado para realização dos demais experimentos propostos neste trabalho. Essa etapa, denominada

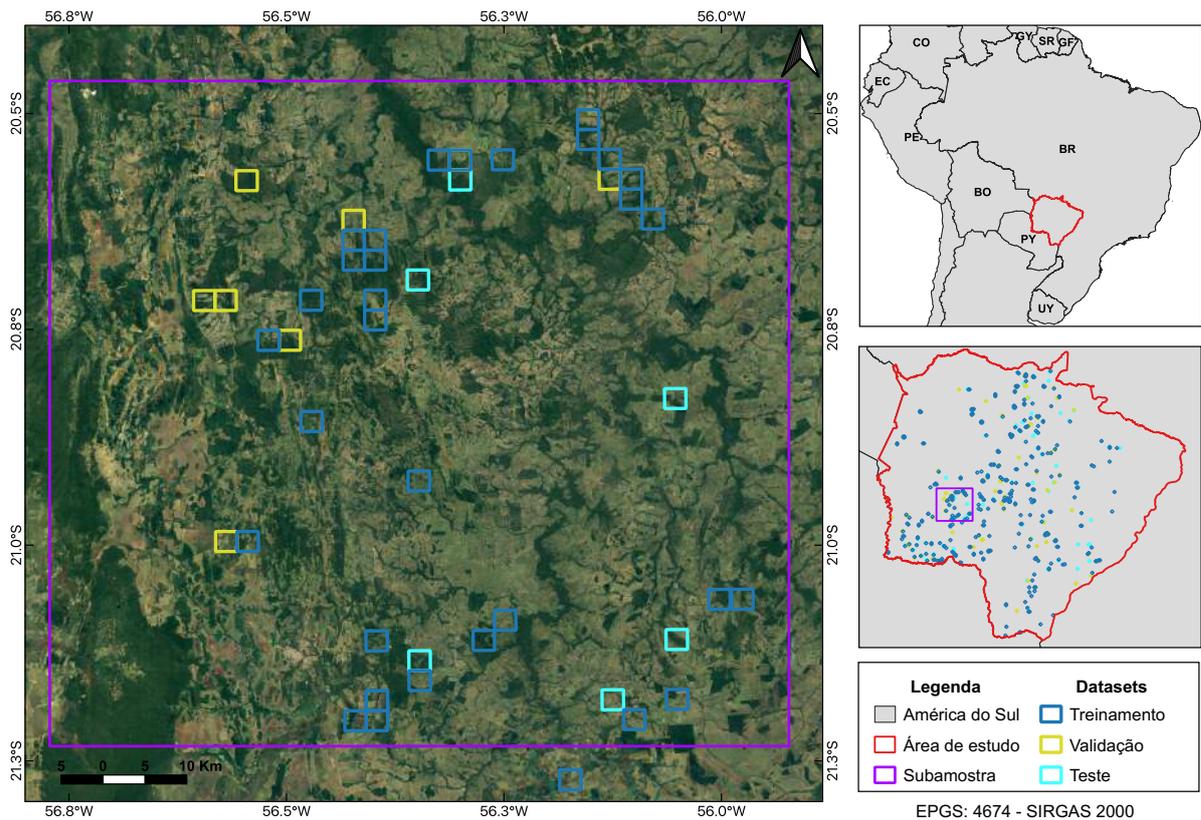


Figura 20 – Conjunto de dados *Sentinel 2*.

“Experimento 1 - Comparação dos conjuntos de dados”, foi realizada com objetivo de analisar e discutir as diferenças dos resultados preliminares obtidos durante o treinamento de redes neurais artificiais nos conjuntos de dados gerados após o processamento de imagens dos satélites Landsat-8 e Sentinel 2.

Posteriormente, foi efetuado o treinamento das redes neurais utilizando as arquiteturas DeepLabv3+, U-Net e Multi-Scale Attention for Semantic Segmentation considerando o melhor *dataset* resultante do experimento 1. Ressalta-se que o foco deste estudo recaiu sobre a abordagem de segmentação semântica, a exemplo do experimento de [Matsak et al. \(2022\)](#), que empregou a U-Net como arquitetura principal. Contudo, é comum o desmatamento ocorrer em regiões de expansão agrícola. Portanto, uma busca na literatura foi realizada na expectativa de encontrar um modelo capaz de capturar o contexto presente nas imagens. Nesse sentido, foi identificada a arquitetura DeepLabV3+, que emprega *atrous separable convolution* para a segmentação semântica. Além disso, considerou-se a viabilidade do uso do mecanismo de atenção, o qual pode destacar as regiões afetadas pelo desmatamento. Estudos sobre a arquitetura *Multi-scale attention* têm demonstrado resultados promissores na literatura sobre segmentação semântica.

Para cada arquitetura foram realizados ajustes de hiperparâmetros e experimentos com diferentes opções de otimizadores empregando a função de perda Binary Cross-

Entropy. Após a conclusão da etapa de treinamento, cada arquitetura foi executada utilizando o conjunto separado para testes, para obtenção dos resultados. A Figura 21 ilustra o processo descrito para mapeamento de desmatamento utilizando a arquitetura U-Net, enfatizando a entrada e saída esperadas.

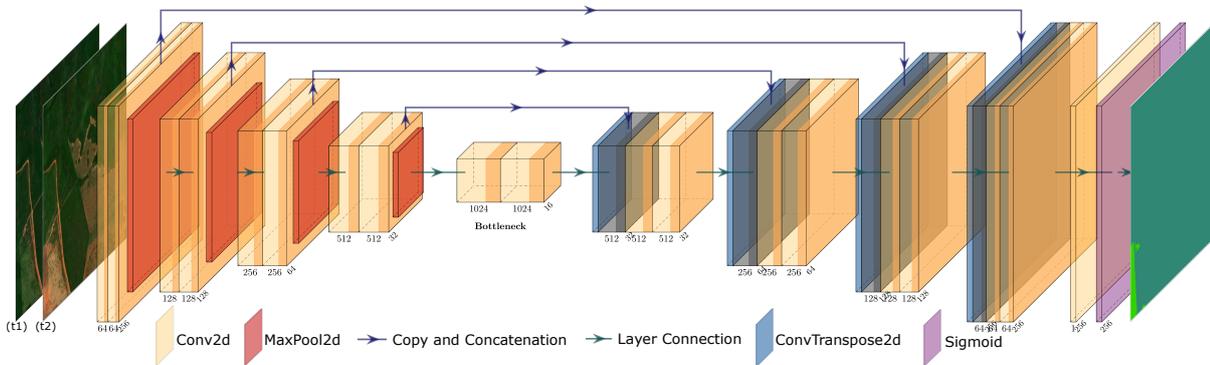


Figura 21 – Representação visual do processo de mapeamento do desmatamento. Do lado esquerdo: um par de imagens adquiridas em datas diferentes, designadas como (t1) e (t2), concatenadas para uma entrada de único fluxo (*single-stream*), também conhecida como *early fusion*. No centro: a arquitetura tradicional U-Net aplicada como *baseline*. Do lado direito: a saída da segmentação semântica, com cada pixel classificado como 1 para área desmatada ou 0 para área não desmatada, levando em consideração apenas a supressão da vegetação nativa que ocorreu no período entre as duas datas.

(Figura produzida utilizando a ferramenta PlotNeuralNet (IQBAL, 2018) baseado no artigo (SOLÓRZANO et al., 2023))

Essa configuração foi chamada de “Experimento 2 - Arquitetura da rede” e buscou analisar a diferença entre os resultados da aplicação de uma arquitetura U-Net tradicional, como *baseline*, com os resultados da aplicação das outras arquiteturas avaliadas neste estudo. Para mitigar o problema de desequilíbrio do conjunto de dados, após a identificação da arquitetura ideal, em comparação com o resultado obtido empregando-se a função de perda *Binary Cross-Entropy*, foram realizados experimentos utilizando a *Focal Loss* e a *Weighted Binary Cross-Entropy*, caracterizadas como funções de perda baseadas em distribuição, que mostraram resultados eficazes para segmentação semântica (JADON, 2020), chamado de “Experimento 3 - Balanceamento do conjunto de dados”.

Em seguida, considerando que a arquitetura com o melhor resultado apresentou uma considerável variação na distribuição entre a métrica *presicion* e a métrica *recall*, foram executados experimentos para calibração do classificador utilizando a técnica de Regressão Isotônica (*Isotonic Regression*) (FAWCETT; NICULESCU-MIZIL, 2007), na expectativa de reduzir os falsos positivos e aumentar o *recall*, denominada como “Experimento 4 - Calibração do classificador”.

Ademais, a alta quantidade de exemplos para classe sem desmatamento pode dificultar a avaliação dos resultados, de forma que a aplicação exclusiva da métrica de acurácia mostraria uma taxa de acerto muito alta, mesmo que as redes considerassem todos os exemplos como não desmatado. Desta forma, as métricas de avaliação F1 Score e IoU foram utilizadas em conjunto com a inspeção visual do resultado da aplicação do melhor

---

modelo no conjunto de testes para proporcionar uma análise apropriada dos resultados, denominando-se este item como “Inspeção visual e análise de resultados”.

Por fim, após a realização dos experimentos supracitados, que expressam os objetivos do presente trabalho, o modelo escolhido como a melhor opção foi disponibilizado ao Ministério Público Estadual, que procederá a sua implantação e implementação de procedimentos para validação dos resultados, bem como aprimoramento da amostra e evolução do modelo.



## 5 Resultados e Discussão

Neste capítulo serão apresentadas as configurações experimentais utilizadas e o resultado dos experimentos executados com intuito de responder as questões levantadas nos objetivos deste trabalho, com ênfase na análise da efetividade dos resultados das arquiteturas DeepLabv3+, U-Net e a Multi-Scale Attention for Semantic Segmentation, que foram aplicadas para detecção de desmatamento no Estado de Mato Grosso do Sul, especialmente nos biomas Cerrado e Mata Atlântica.

### 5.1 Experimento 1 - Comparação dos conjuntos de dados

Ao executar e comparar os resultados de treinamentos preliminares da rede neural artificial U-Net, definida como *baseline*, nos conjuntos de dados gerados após o processamento de imagens dos satélites Landsat-8 e Sentinel 2, obteve-se os resultados apresentados na Tabela 3. Esses evidenciaram que o modelo treinado com o *dataset* gerado a partir das imagens de satélite Sentinel-2 apresentou resultados mais favoráveis.

Tabela 3 – Comparação entre resultados dos conjuntos de dados Lansat-8 e Sentinel 2.

Dataset	Recortes	Arquitetura	Pixel accuracy	F1-Score	IoU
Landsat-8	340	U-Net	0.9771	0.6495	0.4958
Sentinel 2	514	U-Net	<b>0.9822</b>	<b>0.7869</b>	<b>0.6487</b>

Ademais, é relevante salientar que além dos resultados obtidos nos treinamentos preliminares, a observação da qualidade dos recortes resultantes em cada cenário também é importante, considerando que a análise das características de cada metodologia utilizada para definição dos conjuntos de dados foi apresentada como finalidade deste trabalho.

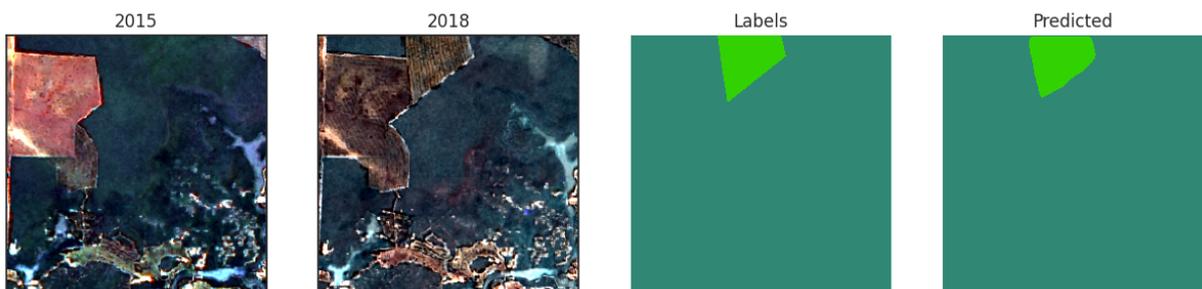


Figura 22 – Exemplo de resultado visualmente promissor utilizando o conjunto de dados Landsat-8.

Desta forma, apesar de algumas das saídas geradas pelo modelo que utilizou o primeiro conjunto de dados parecerem promissoras, como observa-se no exemplo apresentado

na Figura 22, foram encontrados muitos recortes com defeitos, inclusive no conjunto de dados utilizado para testes, conforme demonstrado no exemplo da Figura 23, que possivelmente influenciaram no baixo índice de assertividade resultante do *dataset* com imagens Landsat-8, associados ao menor número de exemplos e maior quantidade de ruído, se comparados ao *dataset* com imagens Sentinel-2.

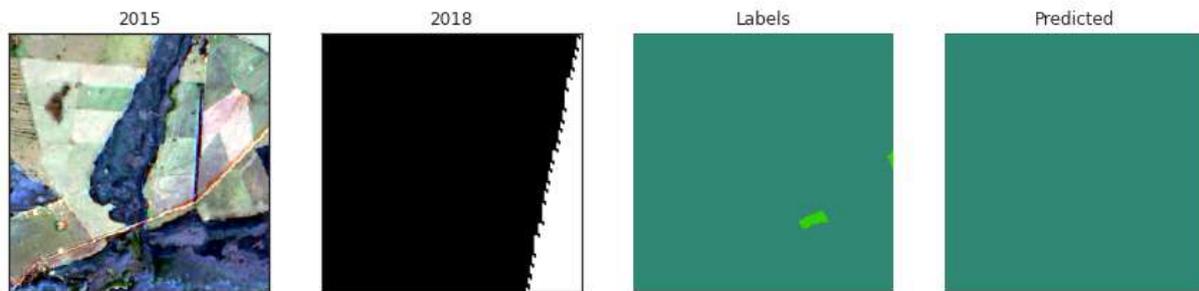


Figura 23 – Exemplo de resultado do conjunto de dados Landsat-8 com defeito.

Assim sendo, evidencia-se que a aplicação de curadoria por especialistas para redução na quantidade de ruído existente no *dataset* é fundamental. No entanto, esta técnica tende a reduzir a quantidade de exemplos existentes na amostra em questão, podendo resultar em um conjunto muito pequeno. Desta forma, faz-se necessário muitos polígonos definidos nos arquivos em formato *Shapefile* e conseqüentemente, a automatização do maior número de etapas para geração dos recortes e da imagem *raster* de saída, com intuito de produzir a maior quantidade de exemplos possíveis e garantir a preservação de uma amostra representativa após a aplicação da curadoria.

A metodologia utilizada na confecção da amostra com dados do satélite Sentinel 2 foi criada de maneira automatizada e passou pela curadoria de especialistas, produzindo exemplos com maior qualidade e obteve os melhores resultados nos treinamentos preliminares das redes neurais, mostrando-se mais adequada para a finalidade deste trabalho. Por esta razão, o segundo *dataset* foi empregado nos demais experimentos.

## 5.2 Experimento 2 - Arquitetura da rede

O resultado do treinamento com arquitetura U-Net tradicional, aplicada como *baseline*, foi comparado com os resultados da aplicação das arquiteturas: DeepLabv3+ e Multi-Scale Attention for Semantic Segmentation, considerando a relevância de seus resultados em trabalhos semelhantes, para atividade de segmentação semântica.

A mesma dimensão de entrada foi utilizada para o treinamento dos modelos, assim como as mesmas métricas de avaliação. Para todos os experimentos foi utilizado o conjunto de dados com recortes do satélite Sentinel 2 com fusão precoce das imagens de entrada. Desta forma, a tecnologia utilizada pode ser classificada como *single-stream* com a mesma configuração de entrada de 256x256x10, sendo 4 canais de cada imagem do satélite Sentinel

Tabela 4 – Comparação entre resultados das arquiteturas U-net, DeepLabv3+ and Multi-Scale Attention segmentados por classe utilizando Binary Cross-Entropy como função de perda. Os valores em negrito representam o maior valor obtido para a classe desmatada.

Architecture	Class	Precision	Recall	F1
U-Net	Deforested (1)	0.8075	0.7674	0.7869
	Non-deforested (0)	0.9896	0.9918	0.9907
DeepLabv3+	Deforested (1)	<b>0.8539</b>	0.7512	0.7993
	Non-deforested (0)	0.9889	0.9943	0.9916
Multi-Scale Attention	Deforested (1)	0.8062	<b>0.7932</b>	<b>0.7997</b>
	Non-deforested (0)	0.9908	0.9915	0.9911

2, quais sejam: *Red*, *Blue*, *Green* e *Infrared* acrescidos de uma camada NDVI, também para cada imagem.

A tabela 4 apresenta o resultado dos modelos mencionados, com ênfase para a *Multi-Scale Attention for Semantic Segmentation* que demonstrou o melhor desempenho em relação a métrica F1-Score. A figura 24 apresenta a curva de perda durante o treinamento desta rede, com a exibição do resultado da função de perda para os conjuntos de treinamento e validação para cada época.

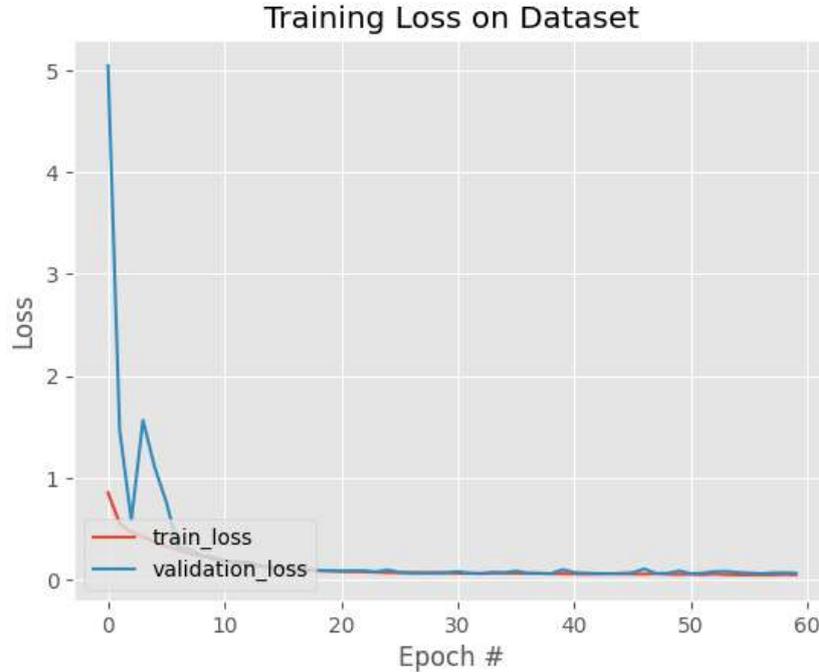


Figura 24 – Curva de perda durante o treinamento da arquitetura *Multi-Scale Attention for Semantic Segmentation*.

Considerando a natureza desbalanceada do conjunto de dados, foram apresentadas a Curva ROC e a AUC ROC na Figura 25. Observa-se que as curvas e a AUC são bastante semelhantes nos resultados apresentados.

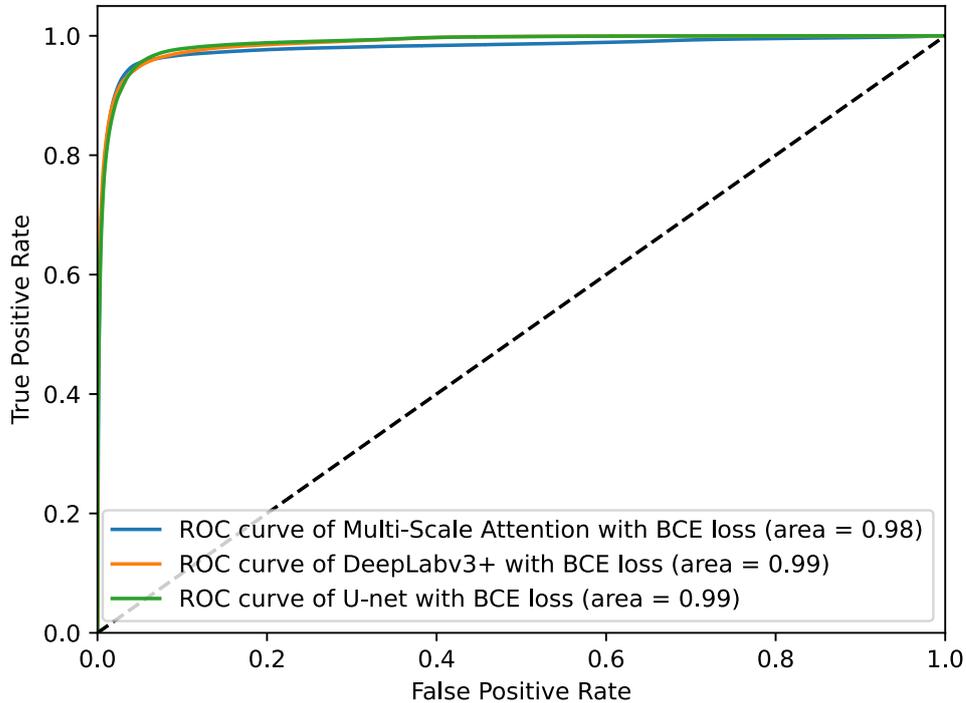


Figura 25 – Curva ROC das arquiteturas U-net, DeepLabv3+ e Multi-Scale Attention.

### 5.3 Experimento 3 - Balanceamento do conjunto de dados

O objetivo deste experimento foi determinar a melhor solução para diminuir o impacto do desbalanceamento do conjunto de dados no resultado dos modelos treinados. Para [Buda, Maki e Mazurowski \(2018\)](#), o problema de bases de dados desbalanceadas é muito comum em aplicações que utilizam aprendizado profundo. Para tanto, algumas técnicas encontradas na literatura podem ser aplicadas para diminuir o impacto deste fator nos resultados.

No caso da segmentação semântica, foi necessário analisar a porcentagem de cobertura de área desmatada em cada imagem para realizar a divisão apropriada das amostras entre os conjuntos de treinamento, validação e testes. Ainda, foi necessário desconsiderar qualquer imagem que não tivesse *pixels* com área desmatada, conforme destacado na metodologia.

Uma possível opção para redução do impacto do balanceamento do conjunto de dados seria a utilização de uma função de perda ponderada, determinando diferentes pesos para cada classe. [Jadon \(2020\)](#) comparou uma considerável quantidade de funções de perda quando utilizadas em atividades de segmentação semântica e observou ótimos resultados para algumas opções como a *Focal Loss* e a *Weighted Binary Cross-Entropy*, que neste experimento foram utilizadas em comparação com a *Binary-Cross-Entropy*, por ter sido a função utilizada para obter o melhor resultado no experimento anterior.

Tabela 5 – Impacto da utilização de diferentes funções de perda utilizando a arquitetura Multi-scale Attention. Os valores em negrito representam o maior valor obtido para a classe desmatada.

Loss	Class	Precision	Recall	F1-Score
Binary	Deforested (1)	0.8062	0.7932	0.7997
Cross-Entropy	Non-deforested (0)	0.9908	0.9915	0.9911
Weighted Binary	Deforested (1)	0.7864	<b>0.8819</b>	<b>0.8314</b>
Cross-Entropy	Non-deforested (0)	0.9947	0.9893	0.9920
Focal Loss	Deforested (1)	<b>0.8481</b>	0.7621	0.8028
	Non-deforested (0)	0.9894	0.9939	0.9917

Na tabela 5 é demonstrado o efeito de utilização de diferentes funções de perda utilizando a melhor arquitetura observada no experimento anterior, de forma que a função de perda que obteve o melhor desempenho foi *Weighted Binary Cross-Entropy*. Esta função de perda, quando configurada para considerar o peso da classe positiva (com desmate) aumentado em 3 vezes em relação ao peso da classe negativa (sem desmate), desempenhou um resultado consideravelmente superior, enquanto a *Focal Loss* não demonstrou aparente ganho de desempenho.

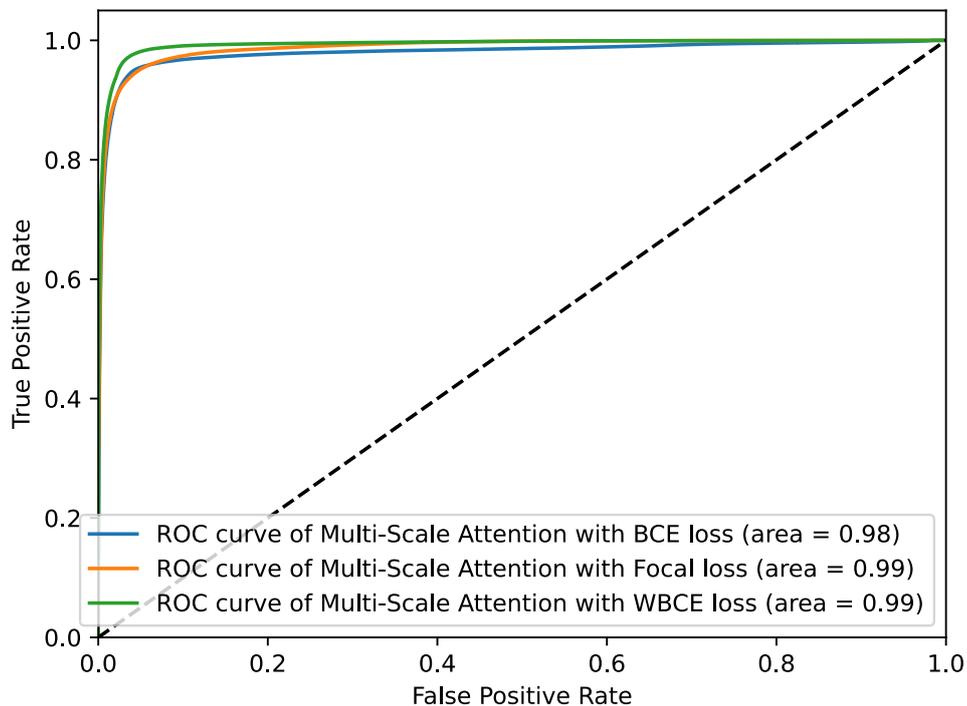


Figura 26 – Curva ROC da arquitetura Multi-scale attention com diferentes funções de perda.

Na Figura 26 são apresentadas a Curva ROC e a AUC ROC para comparação das funções de perda. Observa-se, com foco especial para a linha verde apresentada no canto superior esquerdo, região considerada ideal da ROC, que a curva demonstra um resultado superior para a função de perda *Weighted Binary Cross-Entropy*. Este resultado corrobora

com o resultado apresentado na Tabela 5, que indica que WBCE apresenta um f1-score mais elevado.

## 5.4 Experimento 4 - Calibração do classificador

O *trade-off* entre precision e recall pode ser ajustado utilizando qualquer técnica de calibração do classificador (*score calibration*). Neste trabalho foi aplicada a técnica mais simples de *Isotonic Regression* (FAWCETT; NICULESCU-MIZIL, 2007) para a arquitetura Multi-Scale Attention com a função de perda WBCE. A Tabela 6 apresenta os resultados dessa calibração com o objetivo de reduzir os falsos positivos e aumentar o recall.

Tabela 6 – Resultado do modelo Multi-Scale Attention com a função de perda Weighted Binary Cross-Entropy, aplicando Regressão Isotônica.

Isotonic Regression	Class	Precision	Recall	F1
Não	Deforested (1)	0.7864	0.8819	0.8314
	Non-deforested (0)	0.9947	0.9893	0.9920
Sim	Deforested (1)	0.8453	0.8301	0.8376
	Non-deforested (0)	0.9924	0.9932	0.9928

Este experimento permitiu alcançar a melhor versão do modelo encontrado durante a pesquisa, quando combinados a arquitetura *Multi-Scale Attention* com a função de perda *Weighted Binary Cross-Entropy* e a aplicação de *Isotonic Regression*, chegando a um f1-score de 0.8376 e IoU de 0.7206.

Entre os estudos dedicados à detecção de desmatamento no bioma Cerrado, utilizando imagens do satélite Sentinel-2, foi identificado um conjunto de dados bastante semelhante ao do presente estudo no trabalho de Matosak et al. (2022) (Abordagem 3). Nesse estudo, foi realizada uma análise no estado de Mato Grosso, onde o maior F1-score obtido entre as abordagens aplicadas para esse estado utilizando imagens do satélite Sentinel-2 foi de 0,5428 (MATOSAK et al., 2022). Assim, destaca-se que o resultado alcançado de 0,8376 em F1-score pode ser considerado uma notável contribuição.

## 5.5 Inspeção visual e análise de resultados

Tendo em vista que o principal objetivo deste trabalho foi disponibilizar uma forma automatizada para detecção de desmatamento para o MPMS, foi necessário comparar os resultados obtidos pelas arquiteturas de redes neurais avaliadas, com os resultados atualmente obtidos pelos técnicos, a partir da análise de imagens de satélite de forma manual, para avaliar se a semelhança entre os resultados obtidos foi satisfatória.

No entanto, determinar se a semelhança entre os resultados é apropriada foi um grande desafio, visto que algumas diferenciações poderão ser consideradas significativas pela equipe técnica, tornando necessário corrigir manualmente as imagens processadas pela rede neural artificial, acarretando um possível aumento em sua demanda de trabalho, podendo inviabilizar a implantação ou a utilização deste projeto.

Conforme esperado, foi possível perceber durante a análise do primeiro experimento que, mesmo quando a rede obtém um f1-score muito baixo, a métrica de avaliação de acurácia de *pixels* ainda mostra um valor relativamente alto. Isso acontece devido ao desbalanceamento do conjunto de dados que, ao analisar apenas a quantidade de *pixels* que a rede classificou corretamente sofre muita influência da classe sem desmate por ter a quantidade muito maior de exemplo, em outras palavras, mesmo que a rede classifique todos os *pixels* de todas as imagens como sem desmate, ainda teria uma assertividade alta considerando apenas esta métrica de avaliação. Desta forma, para a finalidade deste trabalho e definição do modelo ideal não foi considerado o resultado da métrica de avaliação *Pixel Accuracy*.

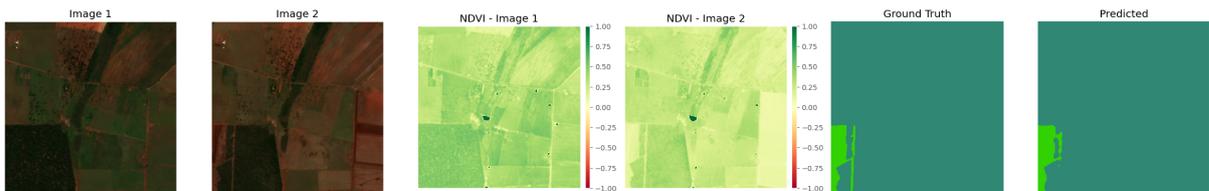


Figura 27 – Exemplo de resultado 1 - Arquitetura Multi-Scale Attention com a função de perda Weighted Binary Cross-Entropy.

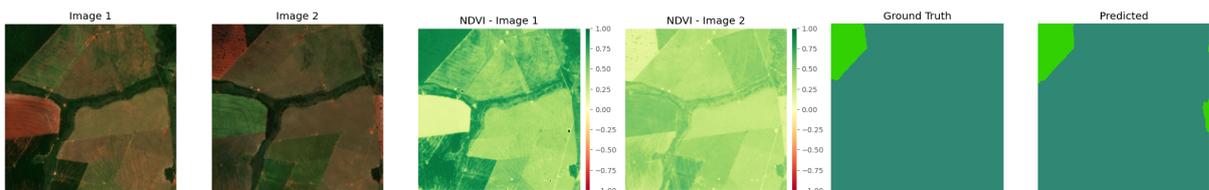


Figura 28 – Exemplo de resultado 2 - Arquitetura Multi-Scale Attention com a função de perda Weighted Binary Cross-Entropy.

No que diz respeito ao f1-score, esta métrica é amplamente reconhecida na literatura como uma ferramenta fundamental para a comparação de resultados. Ao analisar visualmente os resultados dos modelos e compará-los com os resultados obtidos por essa métrica, foi possível constatar uma correlação positiva, gerando confiança nos resultados. Da mesma forma, o *Intersection over Union* (IoU) também foi considerado, e os resultados de ambas as métricas desempenharam um papel crucial na seleção do melhor modelo.

A representação visual do resultado do modelo que apresentou o melhor desempenho pode ser observada nas Figuras 27, 29, 28 e 30. De modo geral, os resultados são bastante impressionantes. Na Figura 27, o modelo detectou uma geometria muito complexa, mas falhou em detectar a estrada. Ademais, ao visualizar o lado direito da figura

é possível observar que houve alteração em uma área de pastagem e a rede identificou a mudança apropriadamente como não sendo desmate. Na Figura 28, o modelo identificou a imagem convexa, porém, detectou alguns falsos positivos na borda direita da imagem. No entanto, ao examinar o lado direito dessa imagem, pode-se perceber uma potencial supressão de vegetação. Determinar se essa vegetação removida deve ser classificada como nativa, caracterizando assim o desmatamento, é uma tarefa complexa. Em tais casos, pode ser interessante o modelo destacar a área para análise adicional por um especialista humano. Na Figura 29, o modelo detectou novamente o desmatamento em uma geometria complexa de forma razoavelmente boa. Por fim, na Figura 30, novamente foram gerados alguns falsos positivos na borda. No entanto, ao observar a imagem real, percebemos que ela apresenta outro cenário complexo, podendo ser considerado ruído no conjunto de testes. Esses resultados foram apresentados ao órgão demandante, que os considerou de qualidade apropriada.

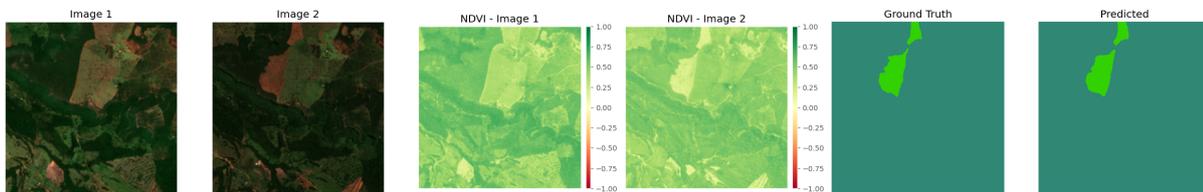


Figura 29 – Exemplo de resultado 3 - Arquitetura Multi-Scale Attention com a função de perda Weighted Binary Cross-Entropy.

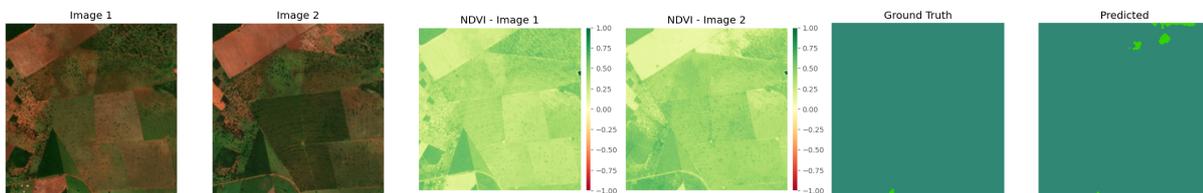


Figura 30 – Exemplo de resultado 4 - Arquitetura Multi-Scale Attention com a função de perda Weighted Binary Cross-Entropy.

## 6 Considerações Finais

A tarefa de mapeamento de desmatamento utilizando séries temporais de imagens de satélite é um desafio que tem sido motivo de diversos experimentos ao redor do mundo. Para solução desta problemática, a utilização de tecnologias como o Aprendizado Profundo para segmentação semântica com fins de detecção de supressões tem se mostrado uma tendência promissora. Contudo, apesar das pesquisas recentes apontarem vantagens na utilização desta tecnologia, ainda pode ser observado uma considerável variação nos resultados divulgados, mesmo em pesquisas que utilizaram biomas, metodologias e arquiteturas similares.

Neste cenário, uma forma de melhorar a análise e comparação dos resultados da aplicação de diferentes técnicas de aprendizado de máquina utilizadas com essa finalidade seria a utilização de bases de dados semelhantes. Contudo, na literatura sobre o tema, a maior parte das publicações não disponibiliza acesso aos dados utilizados, ou na melhor das ocasiões, indica apenas o algoritmo utilizado para download e processamento das imagens, que em uma nova execução provavelmente resultaria em um conjunto distinto.

Neste trabalho, a utilização de um *dataset* com amostras de imagens do satélite Sentinel 2, gerado de maneira automatizada e curado por especialistas, no treinamento de redes neurais profundas resultou em um desempenho significativo para arquitetura *Multi-Scale Attention* com a função de perda *Weighted Binary Cross-Entropy* e a aplicação de Regressão Isotônica, obtendo um f1-score de 0.8376 e IoU de 0.7206.

Para tanto, visando contribuir com a pesquisa e experimentação sobre o tema, o conjunto de dados gerado a partir do processamento das imagens do satélite Sentinel 2 usado no presente trabalho será disponibilizado de maneira gratuita para a comunidade acadêmica, possibilitando a realização de experimentos objetivos e comparações de qualquer técnica desenvolvida recentemente ou posteriormente ao estudo. Ainda que este conjunto não seja considerado perfeito e que certamente ainda permaneça alguma quantidade de ruído em seus dados, o objetivo da disponibilização também visa permitir que novas validações e aprimoramentos sejam efetuados livremente no futuro.

Por fim, pontua-se que o melhor modelo obtido neste trabalho, *Multi-Scale Attention* com 0.8376 de f1-score, foi disponibilizado ao Ministério Público do Estado de Mato Grosso do Sul, que utilizará a tecnologia para auxiliar na garantia de preservação dos ecossistemas encontrados no Estado, especialmente o Cerrado, possibilitando um legado de conservação ambiental e sustentabilidade para as futuras gerações. Ademais, considerando as observações apresentadas no parágrafo anterior, o Órgão também ficará responsável por implementar mecanismos de validação e curadoria para cada resultado

obtido pelo modelo, no sentido de acrescentar os melhores resultados à amostra utilizada para treinamento, aumentando sua capacidade em termos de qualidade e quantidade de dados, para viabilizar a evolução e retreinamento do modelo.

## Referências

- ACHARD, F.; MOLLICONE, D.; STIBIG, H.-J.; AKSENOV, D.; LAESTADIUS, L.; LI, Z.; POPATOV, P.; YAROSHENKO, A. Areas of rapid forest-cover change in boreal eurasia. *Forest Ecology and Management*, Elsevier BV, v. 237, n. 1-3, p. 322–334, dec 2006. Citado na página 20.
- ADARME, M. O.; FEITOSA, R. Q.; HAPP, P. N.; ALMEIDA, C. A. D.; GOMES, A. R. Evaluation of deep learning techniques for deforestation detection in the brazilian amazon and cerrado biomes from remote sensing imagery. *Remote Sensing*, MDPI AG, v. 12, n. 6, p. 910, mar 2020. Citado 3 vezes nas páginas 14, 34 e 35.
- ALSHEHRI, M.; OUADOU, A.; SCOTT, G. J. Deep transformer-based network deforestation detection in the brazilian amazon using sentinel-2 imagery. *IEEE Geoscience and Remote Sensing Letters*, Institute of Electrical and Electronics Engineers (IEEE), p. 1–1, 2024. ISSN 1558-0571. Citado 2 vezes nas páginas 14 e 33.
- ANDRADE, R. B.; COSTA, G. A. O. P.; MOTA, G. L. A.; ORTEGA, M. X.; FEITOSA, R. Q.; SOTO, P. J.; HEIPKE, C. EVALUATION OF SEMANTIC SEGMENTATION METHODS FOR DEFORESTATION DETECTION IN THE AMAZON. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Copernicus GmbH, XLIII-B3-2020, p. 1497–1505, aug 2020. Citado 4 vezes nas páginas 14, 27, 31 e 34.
- ARAGÃO, F. M.; ARAÚJO, F. D. A. da S. Sensoriamento remoto na perícia ambiental do ministério público do estado do piauí: estudo de caso. *Engineering Sciences*, Companhia Brasileira de Producao Cientifica, v. 2, n. 1, p. 17–28, nov 2015. Citado na página 22.
- BEM, P. de; JUNIOR, O. de C.; GUIMARÃES, R. F.; GOMES, R. T. Change detection of deforestation in the brazilian amazon using landsat data and convolutional neural networks. *Remote Sensing*, MDPI AG, v. 12, n. 6, p. 901, mar 2020. Citado 2 vezes nas páginas 14 e 33.
- BORGES, J. P.; LIBÓRIO, M. P.; HADDAD, P. B. Uma nova abordagem para a geovisualização de dados de segurança pública: o caso do ministério público do rio grande do norte. Zenodo, 2019. Citado na página 22.
- BRASIL. Constituição (1988). *Constituição da República Federativa do Brasil*. Brasília, DF: Senado, 1988. Citado 2 vezes nas páginas 13 e 22.
- BRASIL, M. do M. A. *Plano de Ação para Prevenção e Controle do Desmatamento na Amazônia Legal (PPCDAm)*. 2018. Disponível em: [https://combateadesmatamento.mma.gov.br/images/Doc\\_ComissaoExecutiva/Livro-PPCDam-e-PPCerrado\\_20JUN2018.pdf](https://combateadesmatamento.mma.gov.br/images/Doc_ComissaoExecutiva/Livro-PPCDam-e-PPCerrado_20JUN2018.pdf). Acesso em: 02 fevereiro 2024. Citado na página 19.
- BUDA, M.; MAKI, A.; MAZUROWSKI, M. A. A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks*, Elsevier BV, v. 106, p. 249–259, oct 2018. Citado na página 52.

- CHEN, H.; SHI, Z. A spatial-temporal attention-based method and a new dataset for remote sensing image change detection. *Remote Sensing*, MDPI AG, v. 12, n. 10, p. 1662, maio 2020. Disponível em: <<https://doi.org/10.3390/rs12101662>>. Citado na página 31.
- CHEN, L.-C.; YANG, Y.; WANG, J.; XU, W.; YUILLE, A. L. Attention to scale: Scale-aware semantic image segmentation. nov. 2015. Citado 2 vezes nas páginas 24 e 28.
- CHEN, L.-C.; YANG, Y.; WANG, J.; XU, W.; YUILLE, A. L. Attention to scale: Scale-aware semantic image segmentation. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016. p. 3640–3649. Citado na página 28.
- CHEN, L.-C.; ZHU, Y.; PAPANDREOU, G.; SCHROFF, F.; ADAM, H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: *ECCV*. [S.l.: s.n.], 2018. Citado 3 vezes nas páginas 24, 27 e 31.
- COSTA, C. C.; GOMES, L. J.; ALMEIDA, A. P. de. SELEÇÃO DE INDICADORES DE SUSTENTABILIDADE EM FRAGMENTOS FLORESTAIS DE MATA ATLÂNTICA NA BACIA HIDROGRÁFICA DO RIO POXIM-SE POR MEIO DO GEOPROCESSAMENTO. *Revista Eletrônica em Gestão, Educação e Tecnologia Ambiental*, Universidad Federal de Santa Maria, v. 18, n. 1, apr 2014. Citado na página 20.
- CUNHA, N. R. da S.; LIMA, J. E. de; GOMES, M. F. de M.; BRAGA, M. J. A intensidade da exploração agropecuária como indicador da degradação ambiental na região dos cerrados, brasil. *Revista de Economia e Sociologia Rural*, FapUNIFESP (SciELO), v. 46, n. 2, p. 291–323, jun 2008. Citado 3 vezes nas páginas 13, 19 e 20.
- DAUDT, R. C.; SAUX, B. L.; BOULCH, A. *Fully Convolutional Siamese Networks for Change Detection*. 2018. Citado na página 31.
- European Space Agency. *SENTINEL 2: ESA's optical high-resolution mission for GMES operational services*. [S.l.: s.n.], 2012. Citado na página 21.
- FAGUNDES, M. B. B.; GIANETTI, G. W.; OLIVEIRA, D. V. de; DIAS, D. T.; SILVA, L. C. da. Desenvolvimento econômico do estado de mato grosso do sul: Uma análise da composição da balança comercial. *Desenvolvimento em Questão*, Editora Unijui, v. 15, n. 39, p. 112, may 2017. Citado na página 13.
- FAO. *Global forest resources assessment 2015 : how are the world's forests changing*. Rome: Food and Agriculture Organization of the United Nations, 2016. ISBN 9789251092835. Citado na página 19.
- FAO. *The State of the World's Forests 2020*. [S.l.]: FAO and UNEP, 2020. Citado na página 19.
- FAWCETT, T.; NICULESCU-MIZIL, A. Pav and the roc convex hull. *Machine Learning*, Springer, v. 68, p. 97–106, 2007. Citado 2 vezes nas páginas 46 e 54.
- FERREIRA, L. G.; FERREIRA, N. C.; FERREIRA, M. E. Sensoriamento remoto da vegetação: evolução e estado-da-arte. *Acta Scientiarum. Biological Sciences*, Universidade Estadual de Maringá, v. 30, n. 4, nov 2008. Citado na página 20.

- Fundação SOS Mata Atlântica. *Relatório Anual 2019*. 2019. Disponível em: <https://cms.sosma.org.br/wp-content/uploads/2020/09/Relat%C3%B3rio-Anual-SOS-Mata-Atl%C3%A2ntica-2019.pdf>. Acesso em: 26 agosto 2021. Citado 3 vezes nas páginas 13, 19 e 20.
- Fundação SOS Mata Atlântica. *Relatório Anual 2022*. 2022. Disponível em: [https://cms.sosma.org.br/wp-content/uploads/2023/05/SOSMARElatorio-Anual-22\\_dig-pt.pdf](https://cms.sosma.org.br/wp-content/uploads/2023/05/SOSMARElatorio-Anual-22_dig-pt.pdf). Acesso em: 02 fevereiro 2024. Citado na página 13.
- GONÇALVES, A. B.; LOUBET, L.; MOTTA, J.; OLIVEIRA, W.; PINHO, A. Atuação do Ministério Público do Estado de Mato Grosso do Sul na Fiscalização do Desmatamento com Aplicação de Geotecnologias: o programa DNA ambiental. In: \_\_\_\_\_. [S.l.]: Uniedusul Editora, 2021. p. 337–350. ISBN 9786586010626. Citado 4 vezes nas páginas 13, 14, 19 e 23.
- GUO, E.; FU, X.; ZHU, J.; DENG, M.; LIU, Y.; ZHU, Q.; LI, H. *Learning to Measure Change: Fully Convolutional Siamese Metric Networks for Scene Change Detection*. 2018. Citado na página 31.
- HE, K.; ZHANG, X.; REN, S.; SUN, J. *Deep Residual Learning for Image Recognition*. 2015. Citado na página 24.
- IBGE. *Biomass e Sistema Costeiro-Marinho do Brasil*. 2019. Disponível em: <https://www.ibge.gov.br/geociencias/informacoes-ambientais/vegetacao/15842-biomass.html>. Acesso em: 02 fevereiro 2024. Citado 2 vezes nas páginas 13 e 44.
- IM, D. J.; KIM, C. D.; JIANG, H.; MEMISEVIC, R. Generating images with recurrent adversarial networks. fev. 2016. Citado na página 26.
- IMASUL. *Começa em abril inventário de plantas e solo do Cerrado de Mato Grosso do Sul*. 2018. Disponível em: <https://www.imasul.ms.gov.br/comeca-em-abril-inventario-de-plantas-e-solo-do-cerrado-de-mato-grosso-do-sul/>. Acesso em: 26 agosto 2021. Citado na página 20.
- IQBAL, H. *PlotNeuralNet v1.0.0*. 2018. Disponível em: <https://github.com/HarisIqbal88/PlotNeuralNet>. Acesso em: 24 fevereiro 2024. Disponível em: <<https://zenodo.org/record/7261897>>. Citado 2 vezes nas páginas 26 e 46.
- ISAIENKOV, K.; YUSHCHUK, M.; KHRAMTSOV, V.; SELIVERSTOV, O. Deep learning for regular change detection in ukrainian forest ecosystem with sentinel-2. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Institute of Electrical and Electronics Engineers (IEEE), v. 14, p. 364–376, 2021. Disponível em: <<https://doi.org/10.1109/jstars.2020.3034186>>. Citado 2 vezes nas páginas 14 e 31.
- JADON, S. A survey of loss functions for semantic segmentation. *2020 IEEE International Conference on Computational Intelligence in Bioinformatics and Computational Biology*, jun. 2020. Citado 2 vezes nas páginas 46 e 52.
- JENSEN, J. *Sensoriamento remoto do ambiente uma perspectiva em recursos terrestres*. Sao Jose dos Campos: Parentese Editora, 2009. ISBN 856050706X. Citado na página 21.

- KATARKI, G.; RANMALE, H.; BIDARI, I.; CHICKERUR, S. Estimating change detection of forest area using satellite imagery. In: *2019 International Conference on Data Science and Communication (IconDSC)*. IEEE, 2019. Disponível em: <<https://doi.org/10.1109/icondsc.2019.8816956>>. Citado 2 vezes nas páginas 14 e 31.
- KEENAN, R. J.; REAMS, G. A.; ACHARD, F.; FREITAS, J. V. de; GRAINGER, A.; LINDQUIST, E. Dynamics of global forest area: Results from the FAO global forest resources assessment 2015. *Forest Ecology and Management*, Elsevier BV, v. 352, p. 9–20, sep 2015. Citado na página 19.
- KHAN, S. H.; HE, X.; PORIKLI, F.; BENNAMOUN, M. Forest change detection in incomplete satellite images with deep neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, Institute of Electrical and Electronics Engineers (IEEE), v. 55, n. 9, p. 5407–5423, set. 2017. Disponível em: <<https://doi.org/10.1109/tgrs.2017.2707528>>. Citado 2 vezes nas páginas 14 e 31.
- KIYA, H.; NAGAMORI, T.; IMAIZUMI, S.; SHIOTA, S. Privacy-preserving semantic segmentation using vision transformer. *Journal of Imaging*, MDPI AG, v. 8, n. 9, p. 233, ago. 2022. ISSN 2313-433X. Citado na página 23.
- LEE, S.-H.; HAN, K.-J.; LEE, K.; LEE, K.-J.; OH, K.-Y.; LEE, M.-J. Classification of landscape affected by deforestation using high-resolution remote sensing data and deep-learning techniques. *Remote Sensing*, MDPI AG, v. 12, n. 20, p. 3372, out. 2020. Disponível em: <<https://doi.org/10.3390/rs12203372>>. Citado 2 vezes nas páginas 14 e 31.
- LIU, W. T. H. *Aplicações de sensoriamento remoto*. Campo Grande: [s.n.], 2006. Citado na página 21.
- LONG, J.; SHELHAMER, E.; DARRELL, T. Fully convolutional networks for semantic segmentation. nov. 2014. Citado na página 23.
- LUO, X.; LI, X.; WU, Y.; HOU, W.; WANG, M.; JIN, Y.; XU, W. Research on change detection method of high-resolution remote sensing images based on subpixel convolution. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Institute of Electrical and Electronics Engineers (IEEE), v. 14, p. 1447–1457, 2021. Disponível em: <<https://doi.org/10.1109/jstars.2020.3044060>>. Citado na página 31.
- MACHIDA, K.; NAMBU, I.; WADA, Y. Transposed convolution as alternative preprocessor for brain-computer interface using electroencephalogram. *Applied Sciences*, MDPI AG, v. 13, n. 6, p. 3578, mar. 2023. ISSN 2076-3417. Citado na página 26.
- MARETTO, R. V.; FONSECA, L. M. G.; JACOBS, N.; KORTING, T. S.; BENDINI, H. N.; PARENTE, L. L. Spatio-temporal deep learning approach to map deforestation in amazon rainforest. *IEEE Geoscience and Remote Sensing Letters*, Institute of Electrical and Electronics Engineers (IEEE), p. 1–5, 2020. Citado 4 vezes nas páginas 14, 25, 33 e 34.
- MATOSAK, B. M.; FONSECA, L. M. G.; TAQUARY, E. C.; MARETTO, R. V.; BENDINI, H. d. N.; ADAMI, M. Mapping deforestation in cerrado based on hybrid deep learning architecture and medium spatial resolution satellite time series. *Remote*

*Sensing*, MDPI AG, v. 14, n. 1, p. 209, jan. 2022. ISSN 2072-4292. Citado 5 vezes nas páginas 14, 34, 35, 45 e 54.

MENDONÇA, R. A. M. D.; BERNASCONI, P.; SANTOS, R. D.; SCARANELLO, M. Uso das geotecnologias para gestão ambiental. ICV - Instituto Centro de Vida, 2011. Citado na página 22.

MENESES, P. R.; ALMEIDA, T. de. Introdução ao processamento de imagens de sensoriamento remoto. Universidade de Brasília, Brasília, 2012. Citado na página 21.

Microsoft Open Source; MCFARLAND, M.; EMANUELE, R.; MORRIS, D.; AUGSPURGER, T. *Microsoft/PlanetaryComputer: October 2022*. Zenodo, 2022. Disponível em: <<https://zenodo.org/record/7261897>>. Citado na página 41.

MPMS. *Folder NUGEO - Nucleo de Geoprocessamento e Sensoriamento Remoto*. 2011. Disponível em: <https://www.mpms.mp.br/nugeo/folder>. Acesso em: 28 julho 2020. Citado na página 23.

MPMS. *Programa DNA Ambiental aponta 30 mil hectares de área desmatada em todo o Estado*. 2019. Disponível em: <https://www.mpms.mp.br/noticias/2019/11/programa-dna-ambiental-aponta-30-mil-hectares-de-area-desmatada-em-todo-o-estado>. Acesso em: 27 julho 2020. Citado na página 23.

MPMS. *A rapidez na informação é a grande conquista para o enfrentamento ao desmatamento no Estado*. 2020. Disponível em: <https://www.mpms.mp.br/noticias/2020/06/a-rapidez-na-informacao-e-a-grande-conquista-para-o-enfrentamento-ao-desmatamento>. Acesso em: 22 julho 2020. Citado 2 vezes nas páginas 14 e 23.

MYERS, N.; MITTERMEIER, R. A.; MITTERMEIER, C. G.; FONSECA, G. A. B. da; KENT, J. Biodiversity hotspots for conservation priorities. *Nature*, Springer Science and Business Media LLC, v. 403, n. 6772, p. 853–858, feb 2000. Citado 3 vezes nas páginas 13, 19 e 20.

NAZAROVA, T.; MARTIN, P.; GIULIANI, G. Monitoring vegetation change in the presence of high cloud cover with sentinel-2 in a lowland tropical forest region in brazil. *Remote Sensing*, MDPI AG, v. 12, n. 11, p. 1829, jun. 2020. Disponível em: <<https://doi.org/10.3390/rs12111829>>. Citado na página 14.

NOH, H.; HONG, S.; HAN, B. Learning deconvolution network for semantic segmentation. maio 2015. Citado na página 26.

OLLIVEIRA, A. K. M.; MEDEIROS, C. B.; DIETRICH, L. C. ATIVIDADES ANTRÓPICAS e a EFETIVIDADE DA ATUAÇÃO DO MINISTÉRIO PÚBLICO DO ESTADO DE MATO GROSSO DO SUL EM PROPRIEDADES RURAIS. *Revista Gestão & Sustentabilidade Ambiental*, Universidade do Sul de Santa Catarina - UNISUL, v. 7, n. 2, p. 192, abr. 2018. Disponível em: <<https://doi.org/10.19177/rgsa.v7e22018192-205>>. Citado na página 13.

ORTEGA, M. X.; BERMUDEZ, J. D.; HAPP, P. N.; GOMES, A.; FEITOSA, R. Q. EVALUATION OF DEEP LEARNING TECHNIQUES FOR DEFORESTATION DETECTION IN THE AMAZON FOREST. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Copernicus GmbH, IV-2/W7, p. 121–128, sep 2019. Citado 4 vezes nas páginas 14, 31, 39 e 42.

- PAIM, B. L. *Estimativa dos componentes superficiais do balanço hídrico no bioma Mata Atlântica - modelo de superfície INLAND*. Dissertação (Mestrado) — Universidade Federal de Viçosa, 2020. Disponível em: <<https://locus.ufv.br//handle/123456789-27709>>. Citado 2 vezes nas páginas 13 e 19.
- PARANHOS, A. C.; MIOTO, C. L.; PESSI, D. D.; GAMARRA, R. M.; SILVA, N. M. da; RIBEIRO, V. de O.; CHAVES, J. R. *Geotecnologias para aplicações ambientais*. [S.l.]: Uniedusul Editora, 2021. Citado na página 22.
- PARENTE, L.; TAQUARY, E.; SILVA, A.; SOUZA, C.; FERREIRA, L. Next generation mapping: Combining deep learning, cloud computing, and big remote sensing data. *Remote Sensing*, MDPI AG, v. 11, n. 23, p. 2881, dec 2019. Citado 3 vezes nas páginas 14, 34 e 35.
- PARK, S.; LEE, D.; CHOI, Y.; MOON, S. *MultiEarth 2023 Deforestation Challenge – Team FOREVER*. 2023. Citado 2 vezes nas páginas 14 e 33.
- PEDRASSOLI, J. C. *Habitação e sensoriamento remoto: uma análise da expansão urbana na RMSP por meio de imagens de satélite aplicando o modelo linear de mistura espectral*. Tese (Doutorado) — Universidade de Sao Paulo, 2016. Citado na página 21.
- RASLAN, A.; TORRES, T.; MIRANDA, C.; FERNANDES, F.; FILHO, A. P.; ANACHE, J.; ESPÍNDOLA, R.; LIMA, D. V.; BB, O. uso de geotecnologias aplicadas à conservação das áreas de vegetação remanescente e áreas de preservação permanente de hidrografia no município de caracol-ms. *SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO-SBSR*, v. 14, n. 25, p. 4271–78, 2009. Citado na página 22.
- RODRIGUES, J. A. M.; VIOLA, M. R.; ALVARENGA, L. A.; MELLO, C. R.; CHOU, S. C.; OLIVEIRA, V. A.; UDDAMERI, V.; MORAIS, M. A. V. Climate change impacts under representative concentration pathway scenarios on streamflow and droughts of basins in the brazilian cerrado biome. *International Journal of Climatology*, Wiley, v. 40, n. 5, p. 2511–2526, apr 2020. Citado na página 13.
- RONNEBERGER, O.; FISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. maio 2015. Citado 2 vezes nas páginas 24 e 25.
- RUMBLE, H.; ANGEOLETTO, F.; CONNOP, S.; GODDARD, M. A.; NASH, C. Understanding and applying ecological principles in cities. In: *Cities and Nature*. [S.l.]: Springer International Publishing, 2019. p. 217–234. Citado 3 vezes nas páginas 13, 19 e 20.
- SANTOS, J. S. M.; VALERIANO, D. d. M. Análise da paisagem de um corredor ecológico na serra da mantiqueira. *MAIA*, v. 528, p. 7, 2003. Citado na página 22.
- SHELHAMER, E.; LONG, J.; DARRELL, T. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Institute of Electrical and Electronics Engineers (IEEE), v. 39, n. 4, p. 640–651, apr 2017. Citado na página 23.
- SHI, W.; ZHANG, M.; ZHANG, R.; CHEN, S.; ZHAN, Z. Change detection based on artificial intelligence: State-of-the-art and challenges. *Remote Sensing*, MDPI AG, v. 12, n. 10, p. 1688, may 2020. Citado 3 vezes nas páginas 14, 24 e 25.

- SILVA, J. de C.; MONTEIRO, G. O.; PAMBOUKIAN, S. V. D. Introdução ao geoprocessamento. *Congresso Alice Brasil, São Paulo*, p. 155–165, 2016. Citado 2 vezes nas páginas 21 e 22.
- SILVA, J. F.; FARINAS, M. R.; FELFILI, J. M.; KLINK, C. A. Spatial heterogeneity, land use and conservation in the cerrado region of brazil. *Journal of Biogeography*, Wiley, v. 33, n. 3, p. 536–548, mar 2006. Citado na página 13.
- SILVA, J. M. D.; BATES, J. Biogeographic patterns and conservation in the south american cerrado: A tropical savanna hotspot. *Bioscience*, v. 52, p. 225–233, 01 2002. Citado na página 20.
- SOLÓRZANO, J. V.; MAS, J. F.; GALLARDO-CRUZ, J. A.; GAO, Y.; OCA, A. Fernández-Montes de. Deforestation detection using a spatio-temporal deep learning approach with synthetic aperture radar and multispectral images. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier BV, v. 199, p. 87–101, maio 2023. ISSN 0924-2716. Citado 3 vezes nas páginas 14, 31 e 46.
- SONG, K.; JIANG, J. AGCDetNet:an attention-guided network for building change detection in high-resolution remote sensing images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Institute of Electrical and Electronics Engineers (IEEE), v. 14, p. 4816–4831, 2021. Disponível em: <<https://doi.org/10.1109/jstars.2021.3077545>>. Citado na página 31.
- TANCREDI, N. S. H.; ALMEIDA, J. R. de; LINS, G. A.; GUERRA, A. J. T.; JORGE, M. d. C. O. Uso de geotecnologias em laudos periciais ambientais: estudo de caso no município de jacundá, paraíba. *Revista Geografar*, v. 7, n. 1, 2012. Citado na página 22.
- TAO, A.; SAPRA, K.; CATANZARO, B. Hierarchical multi-scale attention for semantic segmentation. maio 2020. Citado 2 vezes nas páginas 24 e 28.
- TETILA, E. C.; TETILA, J. L. C.; PISTORI, H.; SILVA, M. A. B. F. D. Desafios do modelo de desenvolvimento agrícola do estado de mato grosso do sul: uma proposta para o desenvolvimento sustentável. *Interações (Campo Grande)*, Universidade Católica Dom Bosco, p. 615–632, sep 2020. Citado na página 13.
- TORRES, D. L.; TURNES, J. N.; VEGA, P. J. S.; FEITOSA, R. Q.; SILVA, D. E.; JUNIOR, J. M.; ALMEIDA, C. Deforestation detection with fully convolutional networks in the amazon forest from landsat-8 and sentinel-2 images. *Remote Sensing*, v. 13, n. 24, 2021. ISSN 2072-4292. Citado 2 vezes nas páginas 14 e 33.
- U.S. Geological Survey. *Landsat 8 (L8) - Data Users Handbook*. [S.l.: s.n.], 2019. Citado na página 21.
- VACCHIANO, M. C.; SANTOS, J. W.; ANGEOLETTO, F.; SILVA, N. M. Do data support claims that brazil leads the world in environmental preservation? *Environmental Conservation*, Cambridge University Press (CUP), v. 46, n. 2, p. 118–120, nov 2018. Citado 3 vezes nas páginas 13, 19 e 20.
- VIE, J.-C.; HILTON-TAYLOR, C.; STUART, S. *Wildlife in a changing world : an analysis of the 2008 IUCN red list of threatened species*. Gland, Switzerland Barcelona, Spain: IUCN Lynx Edicions, 2009. ISBN 9782831710631. Citado na página 19.

WANG, J.; LIU, B.; XU, K. Semantic segmentation of high-resolution images. *Science China Information Sciences*, Springer Science and Business Media LLC, v. 60, n. 12, nov 2017. Citado na página [23](#).

XU, H.; LIAN, B. Fault detection for multi-source integrated navigation system using fully convolutional neural network. *IET Radar, Sonar & Navigation*, Institution of Engineering and Technology (IET), v. 12, n. 7, p. 774–782, may 2018. Citado na página [24](#).

ZHANG, C.; YUE, P.; TAPETE, D.; JIANG, L.; SHANGGUAN, B.; HUANG, L.; LIU, G. A deeply supervised image fusion network for change detection in high resolution bi-temporal remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier BV, v. 166, p. 183–200, aug 2020. Citado na página [24](#).

ZHAO, F.; SUN, R.; ZHONG, L.; MENG, R.; HUANG, C.; ZENG, X.; WANG, M.; LI, Y.; WANG, Z. Monthly mapping of forest harvesting using dense time series sentinel-1 sar imagery and deep learning. *Remote Sensing of Environment*, Elsevier BV, v. 269, p. 112822, fev. 2022. ISSN 0034-4257. Citado 2 vezes nas páginas [14](#) e [31](#).

ZHAO, H.; SHI, J.; QI, X.; WANG, X.; JIA, J. Pyramid scene parsing network. dez. 2016. Citado 2 vezes nas páginas [24](#) e [27](#).