

INSTITUTO FEDERAL DE SANTA CATARINA

ELIAS ANZINI JUNIOR

Utilização de Redes Neurais Convolucionais para Identificação Automática de Áreas de
Desmatamento na Amazônia

Joinville - SC

09/08/2024

ELIAS ANZINI JUNIOR

Utilização de Redes Neurais Convolucionais para Identificação Automática de Áreas de
Desmatamento na Amazônia

Monografia apresentada ao
Curso de Engenharia Elétrica
do Câmpus Joinville do
Instituto Federal de Santa
Catarina para a obtenção do
diploma de Engenheiro
Eletricista

Orientador: Stefano Romeu
Zeplin

Joinville - SC

09/08/2024

Anzini Junior, Elias.

Utilização de redes neurais convolucionais para identificação automática de áreas de desmatamento na Amazônia/Elias Anzini Junior. – Joinville, SC, 2024. 68 p.

Trabalho de Conclusão de Curso (Graduação) - Instituto Federal de Educação Ciência e Tecnologia de Santa Catarina, Curso de Engenharia Elétrica, Joinville, 2024.

Orientador: Stefano Romeu Zeplin.

1. Redes Neurais Convolucionais (CNN). 2. Desmatamento. Amazonia. I. Instituto Federal de Educação Ciência e Tecnologia de Santa Catarina. II. Título.

Elias Anzini Junior

Utilização de Redes Neurais Convolucionais para Identificação Automática de Áreas de
Desmatamento na Amazônia

Este trabalho foi julgado adequado para obtenção do título em Engenheiro Eletricista, pelo
Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina, e aprovado na sua forma
final pela comissão avaliadora
abaixo indicada.

Joinville, 09 de agosto de 2024.



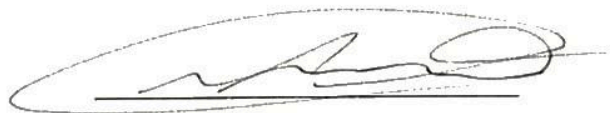
Prof. Stefano Romeu Zeplin, Me.

Orientador

Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina


Prof. Michael Klug Dr.

Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina



Prof. Rodrigo Coral, Dr.

Instituto Federal de Educação, Ciência e Tecnologia de Santa Catarina

AGRADECIMENTOS

Dedico esses agradecimentos acima de tudo a Deus por todos os desafios que colocou em minha vida, à minha família pelo apoio e presença em momentos difíceis, mesmo que de maneira indireta, aos meus amigos de graduação, que considero meu braço direito em minha trajetória, e a todos os professores, que, por mais espinhosos que fossem os caminhos que me fizeram trilhar, sempre me ajudaram com o intuito de me preparar e me tornar um profissional de excelência no futuro.

RESUMO

O desmatamento é uma questão ambiental crítica globalmente que requer monitoramento constante e eficiente, especialmente em regiões como a Amazônia, onde ocorre em larga escala. Com o advento da tecnologia de sensoriamento remoto, a utilização de imagens de satélite tornou-se um método chave para o monitoramento ambiental. No entanto, a eficácia deste método depende fortemente da qualidade e disponibilidade das imagens de satélite, que muitas vezes são prejudicadas por fatores como baixa resolução, cobertura de nuvens e distorções atmosféricas. Atualmente, técnicas de fotointerpretação são empregadas por profissionais do INPE (Instituto Nacional de Pesquisa Espacial) para localização de áreas de desmatamento, entretanto, este processo não é realizado de forma automática. A inteligência artificial tem demonstrado eficácia em uma variedade de campos por meio de técnicas auxiliares, como exemplos: Robótica autônoma, detecção de fraudes, diagnóstico médico assistido por IA, etc. Neste contexto, o presente trabalho de conclusão de curso aborda essa área, com um enfoque particular em redes neurais convolucionais aplicando o modelo U-net para identificação de áreas de desmatamento nas imagens dos satélites Amazonia 1 de baixa resolução espacial e Sentinel-2 de alta resolução espacial. Este modelo foi inicialmente proposto para a análise semântica de imagens médicas, porém ele possui versatilidade e potencial para aplicações além do âmbito médico. O estudo emprega uma metodologia baseada em índices de precisão de segmentação de imagens, tais como Intersecção sobre União (IoU), Precisão, F-Score e Recall. Estas métricas são essenciais para analisar a precisão com que métodos de processamento de imagens conseguem distinguir áreas desflorestadas de outras formas de cobertura terrestre, dando uma base quantitativa para avaliar a eficácia de diferentes fontes de dados de sensoriamento remoto.

Palavras-Chave: Desmatamento, U-Net, Sentinel-2, Amazonia 1, Segmentação de Imagens.

ABSTRACT

Deforestation is a critical global environmental issue that requires constant and efficient monitoring, especially in regions like the Amazon, where it occurs on a large scale. With the advent of remote sensing technology, the use of satellite imagery has become a key method for environmental monitoring. However, the effectiveness of this method heavily depends on the quality and availability of satellite images, which are often compromised by factors such as low resolution, cloud cover, and atmospheric distortions. Currently, photointerpretation techniques are employed by professionals from INPE (National Institute for Space Research) to locate deforestation areas, however, this process is not automated. Artificial intelligence has proven effective in a variety of fields through auxiliary techniques, such as autonomous robotics, fraud detection, and AI-assisted medical diagnosis, etc. In this context, this thesis addresses this area, with a particular focus on convolutional neural networks applying the U-net model for the identification of deforestation areas in the images from the high spatial resolution Sentinel-2 satellites and the low spatial resolution Amazonia 1. This model was initially proposed for the semantic analysis of medical images, yet it has versatility and potential for applications beyond the medical scope. The study employs a methodology based on image segmentation precision indices, such as Intersection over Union (IoU), Precision, F-Score, and Recall. These metrics are essential for analyzing the accuracy with which image processing methods can distinguish deforested areas from other forms of land cover, providing a quantitative basis to evaluate the effectiveness of different remote sensing data sources.

Keywords: Deforestation, U-Net, Sentinel-2, Amazonia 1, Image Segmentation.

LISTA DE ILUSTRAÇÕES

Figura 1: Taxas de Desmatamento - Amazônia Legal (INPE, 2023).....	14
Figura 2: Mudanças Observadas em Extremos de Calor (IPCC, 2021).....	17
Figura 3: Evolução do Aquecimento da Superfície Terrestre entre 1850 - 2020 (IPCC, 2021)..	17
Figura 4: Contorno da Amazônia segundo o critério político-administrativo. Fonte: (PNUMA et al, 2008).....	19
Figura 5: Contorno da Amazônia Legal. Fonte: (IBGE, 2022).....	20
Figura 6: A taxa de desmatamento da Amazônia brasileira em 2020 é a maior da década. Fonte: (SILVA JUNIOR et al, 2021).....	21
Figura 7: Proporção do desmatamento ilegal nos estados da Amazônia e do Matopiba. Fonte: (VALDIONES, Ana Paula et al, 2021).....	22
Figura 8: Localização das cenas Landsat que recobrem a Amazônia Legal Brasileira. Fonte: (DE ALMEIDA, CLAUDIO APARECIDO et al., 2022).....	23
Figura 9: Aplicação da composição colorida PRODES. Fonte: (DE ALMEIDA, CLAUDIO APARECIDO et al., 2022).....	25
Figura 10: Aplicação da composição colorida e MLME DETER . Fonte: (DE ALMEIDA, CLAUDIO APARECIDO et al., 2022).....	26
Figura 11: Categorias de Desmatamento. Adaptado. Fonte: (DINIZ, CESAR GUERREIRO et al., 2015).....	27
Figura 12: Áreas da Inteligência Artificial. Fonte: (MONTEIRO, ROSANGELA et al, 2022)..	29
Figura 13: Programação Clássica x Machine Learning. Fonte: (GRANDO, 2022).....	30
Figura 14: Modelo Perceptron. Fonte: (OSÓRIO, BITTENCOURT, 2000).....	31
Figura 15: Áreas de Aplicação do Deep Learning. Adaptado. Fonte: (TAYE, 2023).....	32
Figura 16: Exemplo de cálculo de Convolução. Fonte: (OLIVEIRA, 2022).....	33
Figura 17: Exemplo de convolução na imagem. Fonte: (SANTOS, 2021).....	34
Figura 18: Operação de Convolução com Padding. Fonte: (SANTOS, 2021).....	34
Figura 19: Mapas de características em uma convolução. Fonte: (PHIL, 2017).....	35
Figura 20: Operação de Max Pooling. Fonte: (SANTOS, 2021).....	35
Figura 21: Rede Convolutiva. Adaptado. Fonte: (FERREIRA, 2017).....	36
Figura 22: Cálculo de Convolução Transposta. Adaptado. Fonte: (ZHANG., LIPTON, et al, 2023).....	37
Figura 23: Convolução transposta em imagem de tamanho 2x2, filtro 3x3, stride 2. Fonte: (LANE, 2018).....	37
Figura 24: Exemplo de Convolução Dilatada. Fonte: (Chen et al., 2017).....	38
Figura 25: (a) Encoder (b) Convolução Atrous. Fonte: (Chen et al., 2017).....	38
Figura 26: Funções de Ativação. Adaptado. Fonte (CUNHA, 2020).....	39
Figura 27: Funcionamento ReLU. Fonte (CUNHA, 2020).....	39
Figura 28: Segmentação Semântica. Fonte: (SILVA, 2021).....	40
Figura 29: Arquitetura U-NET. Fonte: (VILELA, 2024).....	41
Figura 30: Arquitetura U-NET++. Fonte: (ZHOU, SIDDIQUEE, et al, 2018).....	42

Figura 31: Funcionamento do Atrous Spatial Pyramid Pooling. Fonte: (Chen et al, et al, 2018).....	42
Figura 32: Arquitetura DeepLabV3+. Fonte: (Chen et al, et al, 2018).....	43
Figura 33: Imagens do satélite Amazonia 1. Fonte: INPE (2024).....	44
Figura 34: Plataforma TerraBrasilis. Fonte: Terra Brasilis (2024).....	45
Figura 35: Imagens do satélite Amazonia 1 e shapefile de desmatamento. Fontes: Terra Brasilis (2024), INPE (2024).....	45
Figura 36: Imagens do satélite Amazonia 1 e suas máscaras de desmatamento. Fontes: Terra Brasilis (2024), INPE (2024).....	46
Figura 37: Imagem do satélite Amazonia 1 e sua máscaras de desmatamento. Fonte: Terra Brasilis (2024), INPE (2024).....	46
Figura 38: Imagem do satélite Sentinel-2. Fonte: (BRAGAGNOLO, LUCIMARA, et al, 2021).....	47
Figura 39: Matriz de Confusão Exemplificada. Fonte: (SAMIR, 2024).....	48
Figura 40: Intersecção sobre União. Fonte: (MATOS, 2024).....	49
Figura 41: QGIS. Fonte: (AUTOR, 2024).....	51
Figura 42: Treino e Validação IoU, Acurácia e Loss. Sentinel-2 Fonte: (AUTOR, 2024).....	55
Figura 43: Segmentação Semântica. Sentinel-2 Fonte: (AUTOR, 2024).....	56
Figura 44: Treino e Validação IoU, Acurácia e Loss. Amazonia 1. Fonte: (AUTOR, 2024)...	58
Figura 45: Segmentação Semântica. Amazonia 1. Fonte: (AUTOR, 2024).....	59

LISTA DE TABELAS

Tabela 1: Resultados Sentinel-2. Fonte: (AUTOR, 2024).....	54
Tabela 2: Resultados Sentinel-2. Informações Computacionais. Fonte: (AUTOR, 2024).....	54
Tabela 3: Resultados Amazonia 1. Fonte: (AUTOR, 2024).....	57
Tabela 4: Resultados Amazonia 1. Informações Computacionais. Fonte: (AUTOR, 2024)....	57
Tabela 5: Resultados de trabalhos relacionados (AUTOR, 2024).....	60

LISTA DE ABREVIATURAS E SIGLAS

- ABNT – Associação Brasileira de Normas Técnicas
- ASPP - Atrous Spatial Pyramid Pooling
- CIUP - *Centro de Pesquisa da Universidad del Pacífico*
- CNNs - *Convolutional Neural Networks* - Redes Neurais Convolucionais
- DETER - Sistema de Detecção do Desmatamento em Tempo Real
- GAN - *Generative Adversarial Network* - Redes Adversárias Generativas
- IA - Inteligência Artificial
- IBGE – Instituto Brasileiro de Geografia e Estatística
- IFSC – Instituto Federal de Santa Catarina
- INPE - Instituto Nacional de Pesquisas Espaciais
- IOU - *Intersection over Union* - Intersecção sobre União
- MLME - Modelo Linear de Mistura Espectral
- MLP - *Multiplayer Perceptron*
- MSI - *MultiSpectral Instrument* - Instrumento Multiespectral
- NIR - *Near Infrared* - Infravermelho Próximo
- OTCA - Organização do Tratado de Cooperação Amazônica
- PPCDAm - Plano de Ação para Prevenção e Controle do Desmatamento na Amazônia Legal
- PNUMA - Programa das Nações Unidas para o Meio Ambiente
- PRODES - Programa de Cálculo do Desflorestamento da Amazônia
- ReLU - *Rectified Linear Unit* - Unidade Linear Retificada
- RGB - *Red, Green, Blue* - Vermelho, Verde, Azul

SUMÁRIO

1 INTRODUÇÃO.....	14
1.1 OBJETIVOS.....	16
1.1.1 OBJETIVO GERAL.....	16
1.1.2 OBJETIVOS ESPECÍFICOS.....	16
1.1.2 JUSTIFICATIVA DO TRABALHO.....	16
2 DESENVOLVIMENTO.....	19
2.1 REVISÃO DA LITERATURA.....	19
2.1.1 AMAZÔNIA INTERNACIONAL E AMAZÔNIA LEGAL.....	19
2.1.2 DESMATAMENTO NA AMAZÔNIA.....	20
2.1.3 PROGRAMAS DE PREVENÇÃO AO DESMATAMENTO.....	22
2.1.3.1 PRODES.....	22
2.1.3.2 DETER.....	24
2.1.3.3 TÉCNICAS DE CLASSIFICAÇÃO DO DESMATAMENTO.....	24
2.1.3.4 CLASSIFICAÇÃO DO DESMATAMENTO.....	27
2.1.4 SENSORIAMENTO REMOTO.....	28
2.1.4.1 SENTINEL 2.....	28
2.1.4.2 AMAZONIA 1.....	28
2.1.5 INTELIGÊNCIA ARTIFICIAL.....	29
2.1.6 APRENDIZADO DE MÁQUINA.....	30
2.1.7 REDES NEURAIS ARTIFICIAIS.....	31
2.1.8 APRENDIZAGEM PROFUNDA.....	32
2.1.9 REDES NEURAIS CONVOLUCIONAIS.....	33
2.1.9.1 CAMADA CONVOLUCIONAL.....	33
2.1.9.2 CAMADA DE AGRUPAMENTO/REDUÇÃO.....	35
2.1.9.3 CAMADA FINAL.....	36
2.1.9.4 CONVOLUÇÃO TRANSPOSTA.....	36
2.1.9.5 CONVOLUÇÕES DILATADAS.....	37
2.1.9.6 FUNÇÕES DE ATIVAÇÃO.....	38
2.1.10 SEGMENTAÇÃO SEMÂNTICA.....	39
2.1.10.1 ARQUITETURA U-NET.....	40
2.1.10.2 ARQUITETURA U-NET ++.....	41
2.1.10.3 DEEPLAB V3 +.....	42
3 METODOLOGIA.....	44
3.1 COLETA DE DADOS.....	44
3.1.1 AQUISIÇÃO DE IMAGENS AMAZONIA 1.....	44
3.1.2 AQUISIÇÃO DE IMAGENS SENTINEL-2.....	47
3.2 AJUSTE E PREPARAÇÃO DOS DADOS.....	47
3.3 DESENVOLVIMENTO DO MODELO DE SEGMENTAÇÃO SEMÂNTICA.....	48
3.4 AVALIAÇÃO DE DESEMPENHO.....	48

3.4.1 IOU.....	48
3.4.2 ACURÁCIA.....	49
3.4.3 RECALL.....	49
3.4.4 F1-SCORE.....	50
3.5 RECURSOS UTILIZADOS PARA DESENVOLVIMENTO.....	50
3.5.1 GOOGLE COLAB.....	50
3.5.2 TENSOR FLOW E KERAS.....	50
3.5.3 GDAL.....	51
3.5.4 QGIS.....	51
4.1 PARÂMETROS UTILIZADOS.....	52
4.1.1 TAMANHO DO LOTE.....	52
4.1.2 OTIMIZADORES.....	52
4.1.3 INICIALIZADORES.....	52
4.1.4 TAXA DE APRENDIZADO.....	53
4.1.5 ÉPOCAS E PARADA ANTECIPADA.....	53
4.1.6 AUMENTO DE DADOS.....	53
4.2 ANÁLISE DO DESEMPENHO.....	53
5 CONCLUSÃO.....	61

1 INTRODUÇÃO

A Amazônia é atualmente um dos principais biomas tanto para o Brasil quanto para o mundo. Toda sua biodiversidade e recursos naturais disponíveis a tornam foco de interesse para governos e empresas de diversos setores, como exemplos: madeireiro, mineral e agropecuário. A sua crescente exploração econômica, evidenciados pelos níveis de desmatamento entre 2012 a 2021, levantam questões cruciais sobre os impactos negativos ao mundo que o desmatamento pode ocasionar. (DA COSTA PORTELLA, 2022).

Um dos principais problemas enfrentados nessa região é o desmatamento. Conforme dados na figura 1 do Instituto Nacional de Pesquisas Espaciais (INPE, 2023), observa-se uma tendência preocupante de aumento na área desmatada na Amazônia Legal, indo de 4571 km² em 2012 para 13000 km² em 2021. Esta situação reflete não apenas a perda de cobertura florestal, mas também repercute diretamente na biodiversidade, no equilíbrio climático e nas comunidades locais.

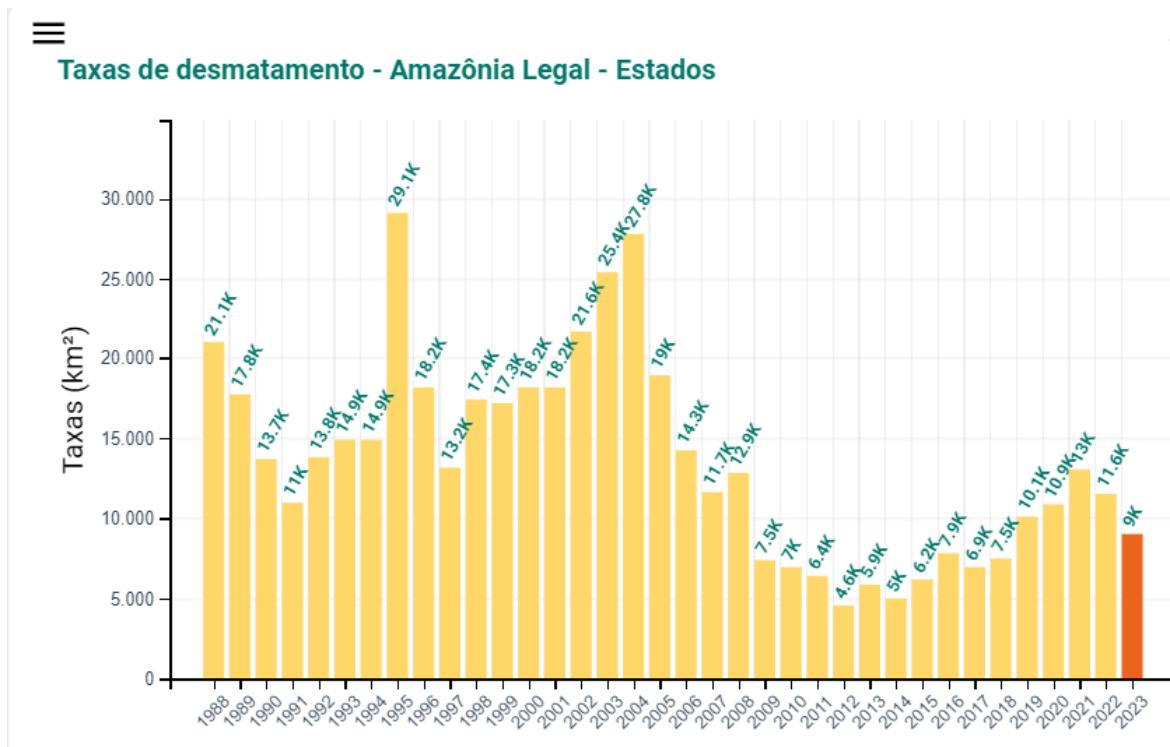


Figura 1: Taxas de Desmatamento - Amazônia Legal (INPE, 2023)

A Amazônia é a maior floresta do mundo, e possui um papel imprescindível para os regimes de chuva e temperaturas principalmente na América do Sul. As interações entre a floresta e atmosfera influenciam os padrões climáticos e hídricos de todo o planeta, como

tem-se observado nos diferentes eventos climáticos no mundo todo. O desmatamento, portanto, pode ter implicações significativas na alteração desses padrões, elevando fenômenos como mudanças climáticas. Um exemplo notório do que a humanidade precisa impedir, é outra devastação como a realizada na Mata Atlântica, que perdeu mais de 90% de sua cobertura original ao longo de cinco séculos (NOBRE AD, 2014).

Devido aos recorrentes problemas relacionados ao desmatamento, a demanda pelo sensoriamento remoto para monitoramento do bioma deve aumentar ao longo dos anos. Conforme descrito por Meneses e Almeida (2012), o sensoriamento remoto é uma ferramenta poderosa e eficiente para a coleta automática de informações sobre recursos terrestres em uma perspectiva global. Desenvolvido nos anos 1960, essa tecnologia evoluiu continuamente com diversos lançamentos de satélites ao espaço com cada vez melhores tecnologias embarcadas, fornecendo dados cruciais para o monitoramento da Terra. Entretanto, seu uso não é perfeito, devido a utilidade dessas imagens muitas vezes ser restringida pela sua resolução, disponibilidade ou qualidade.

Atualmente, a análise do desmatamento na Amazônia é feita pelo INPE, através dos programas PRODES (Projeto de Monitoramento do Desmatamento na Amazônia Legal por Satélite) e DETER (Sistema de Detecção do Desmatamento em Tempo Real). Na metodologia destes programas, as imagens selecionadas são analisadas de forma manual, onde os especialistas utilizam técnicas de fotointerpretação para delimitar os polígonos de desmatamento (INPE, 2022). Na busca de novas soluções para o monitoramento, o avanço da visão computacional e inteligência artificial se mostram promissores. A aplicação de técnicas de *deep learning* (aprendizado profundo), especificamente através de redes neurais convolucionais, pode ser uma abordagem poderosa para interpretar automaticamente grandes volumes de dados provenientes de imagens de satélite.

A identificação precisa e detalhada de áreas de desmatamento é crucial para o controle de danos e a implementação de ações preventivas. Neste contexto, o presente trabalho propõe a utilização de inteligência artificial e redes neurais convolucionais para identificação automática de regiões desmatadas utilizando imagens dos satélites Sentinel-2 e Amazonia 1.

1.1 OBJETIVOS

1.1.1 OBJETIVO GERAL

Realizar a detecção automática das áreas de desmatamento através do uso de imagens de satélite fornecidas pelos satélites Amazonia 1 e Sentinel-2.

1.1.2 OBJETIVOS ESPECÍFICOS

Abaixo são descritos os objetivos específicos do projeto:

- Desenvolver um algoritmo para extração de imagens de satélite e máscaras de desmatamento, capaz de, através das coordenadas fornecidas, inserir a imagem e sua máscara de desmatamento no conjunto de dados.
- Explorar as fundamentações teóricas sobre a inteligência artificial, redes neurais convolucionais e da arquitetura U-Net, e como elas têm sido aplicadas em problemas semelhantes;
- Desenvolver um algoritmo para coleta de imagens dos satélites ou buscar por banco de dados confiáveis para coleta de dados;
- Utilizar arquitetura de segmentação semântica U-Net para detecção de áreas desmatadas.
- Avaliar o desempenho dos modelos de segmentação semântica em termos de Precisão, Intersecção sobre União (*IoU*), *F-Score* e *Recall*;
- Realizar a validação e documentar.

1.1.2 JUSTIFICATIVA DO TRABALHO

O desmatamento no mundo todo tem provocado impactos ambientais cuja consequências tem causado cada vez mais relatos de desastres naturais. Através da figura 2, fornecida no relatório de Mudança do Clima em 2021 pelo Painel Intergovernamental sobre Mudanças Climáticas, é possível observar as mudanças em extremos de calor ocorridas no planeta, onde quase todas as regiões sofreram com esse tipo de alteração climática e a maioria possui alta confiança na contribuição humana para essas mudanças. Além disso, através do mesmo relatório divulgado, figura 3, é importante observar a alteração climática em decorrência do fator humano, o aumento expressivo na temperatura da superfície global no período observado entre 1850 e 2019, que foi estimado em 1,07°C, dentro de uma faixa provável de 0,8°C a 1,3°C (IPCC, 2021).

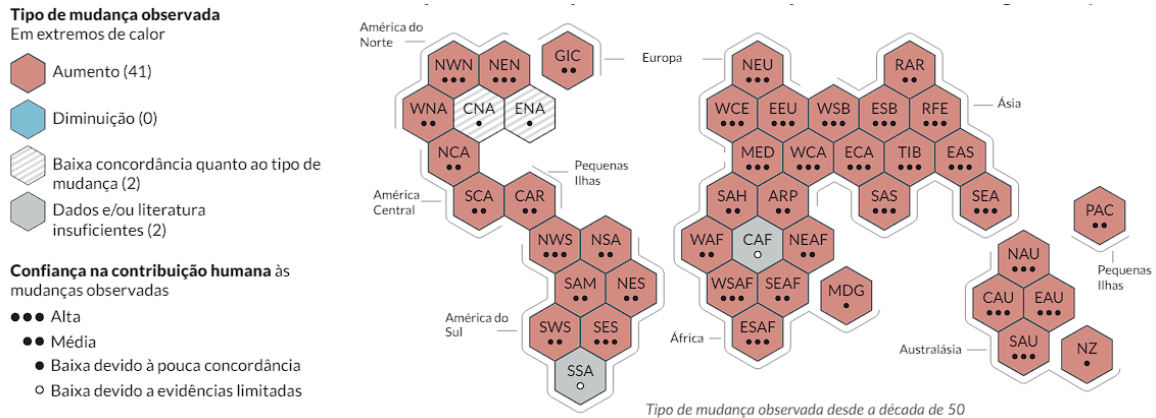


Figura 2: Mudanças Observadas em Extremos de Calor (IPCC, 2021).

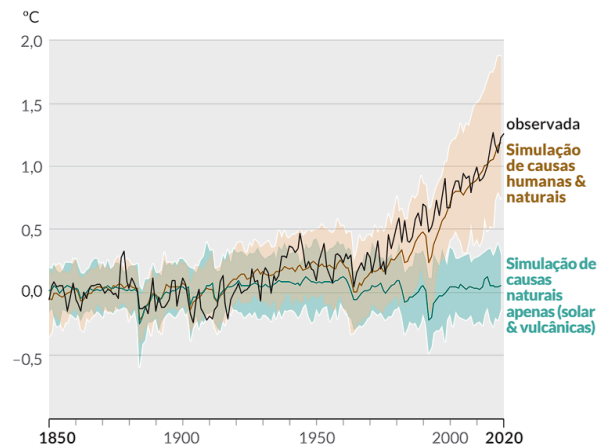


Figura 3: Evolução do Aquecimento da Superfície Terrestre entre 1850 - 2020 (IPCC, 2021).

Nesse contexto, o monitoramento e diminuição do desmatamento é essencial para a redução dos fatores humanos no aquecimento global e degradação ambiental, especialmente em regiões como a Amazônia, que enfrentam taxas significativas de perda florestal. As imagens de satélite têm sido uma ferramenta fundamental nesse esforço, fornecendo uma visão abrangente e atualizada das mudanças na cobertura terrestre. No entanto, a utilidade dessas imagens pode ser limitada por desafios como baixa resolução, cobertura de nuvens, e restrições de acesso, especialmente com satélites de alta resolução. Estes desafios reduzem a eficácia do monitoramento do desmatamento, podendo gerar respostas atrasadas, equivocadas ou incompletas. Deste modo, novas técnicas capazes de detectar regiões de desmatamento através de imagens de baixa qualidade são fundamentais para aprimorar os métodos atuais de monitoramento.

A aquisição de sistemas espaciais para suprir as necessidades de monitoramento de um país representa um investimento significativo. Como ilustração, o Amazonia 1, o primeiro

satélite inteiramente desenvolvido pelo Brasil, lançado em 2021, teve um custo de aproximadamente 380 milhões de reais, oferecendo bandas visíveis com resolução espacial máxima de até 64 metros (Revista Pesquisa Fapesp, 2021). Em comparação, o satélite Landsat 8, um projeto dos Estados Unidos, teve seu custo de lançamento estimado em 855 milhões de dólares, possuindo uma das bandas visíveis com resolução espacial máxima de 30 metros (CBS News, 2013).

Além disso, Magrinelli et al. (2022) destacam que o Reino Unido reconhece a falta de meios espaciais necessários para obter informações ambientais adequadas. Através das informações expostas, fica nítido a necessidade de buscar novas metodologias que consigam extrair informações precisas através de meios menos custosos e com alta disponibilidade de informações.

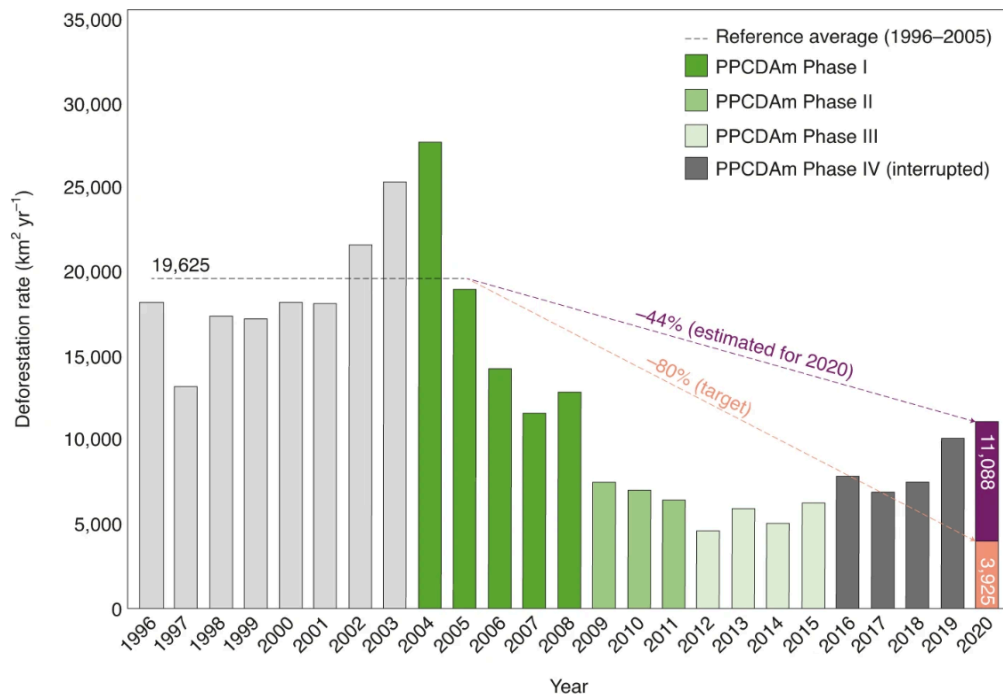


Figura 6: A taxa de desmatamento da Amazônia brasileira em 2020 é a maior da década. Fonte: (SILVA JUNIOR et al, 2021).

Como parte de seus esforços para atingir as metas estabelecidas de redução das emissões de gases de efeito estufa, o Brasil implementou a Política Nacional de Mudanças Climáticas em 2009. Esta política estabeleceu o objetivo ambicioso de diminuir a taxa de desmatamento na Amazônia em 80% até o ano de 2020, limitando o desmatamento máximo anual a 3 925 km². Em 2012, a taxa de desmatamento na região atingiu seu nível mais baixo historicamente, refletindo progressos significativos nesse sentido. No entanto, a partir de 2013, começou-se a observar uma tendência preocupante de aumento no desmatamento. Essa tendência assentou-se, e em 2021, a região registrou o maior nível de desmatamento desde 2006, indicando um retrocesso nos esforços de conservação do país. (SILVA JUNIOR et al, 2021).

O desmatamento pode ser categorizado de diferentes formas, dependendo das causas, métodos e impactos associados. O desmatamento legal ocorre quando a remoção da vegetação é autorizada pelas autoridades competentes, seguindo as leis e regulamentos ambientais do país. Já o desmatamento ilegal, é realizado em contravenção às leis e regulamentos locais, nacionais ou internacionais, muitas vezes associado a práticas corruptas e exploração predatória. Na figura 7, é evidenciado a taxa de desmatamento ilegal nos estados da Amazônia e do Matopiba.

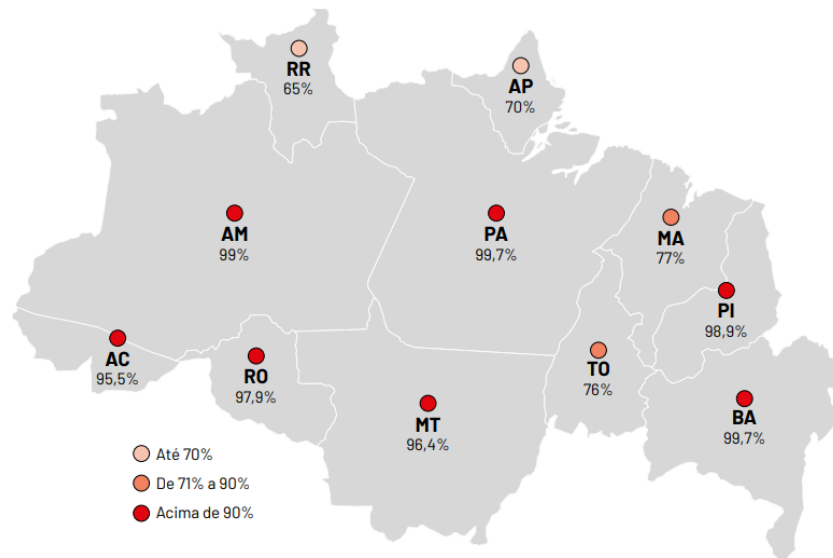


Figura 7: Proporção do desmatamento ilegal nos estados da Amazônia e do Matopiba. Fonte: (VALDIONES, Ana Paula et al, 2021).

2.1.3 PROGRAMAS DE PREVENÇÃO AO DESMATAMENTO

No Brasil existem dois sistemas para monitoramento desenvolvidos e operados pelo INPE que desempenham papéis complementares e cruciais, o PRODES e o DETER.

2.1.3.1 PRODES

O PRODES representa uma ferramenta essencial no monitoramento ambiental da Amazônia. Utilizado desde 1988, ele surgiu em resposta à necessidade de um sistema eficaz para monitorar a extensão do desmatamento na Amazônia. Através dele, todo ano são fornecidos dados referentes ao desmatamento anual. A figura 1 foi extraída do relatório anual do PRODES (INPE, 2023).

Por ser um relatório de alta precisão, o PRODES utiliza imagens de alta qualidade, selecionando as melhores coletadas ao longo do ano para sua elaboração. Essas imagens possuem resolução espacial de 20 a 30 metros e incluem pelo menos três bandas espectrais. O programa utiliza principalmente as imagens dos satélites Landsat-8 e Landsat-9 (EUA), mas também pode empregar imagens auxiliares do SENTINEL-2 (União Europeia) e do CBERS-4/4A (colaboração Brasil-China). Além disso, ele tem somente capacidade de localizar polígonos de desmatamento superiores a 1 hectare, e para manter a série temporal, só

são incluídos no cálculo de desmatamento os polígonos iguais ou superiores a 6,25 hectares (DE ALMEIDA, CLAUDIO APARECIDO et al., 2022).

Na figura 8, é exibido como é realizada a divisão de imagens Landsat para sensoriamento da Amazônia Legal Brasileira. Para isso, são necessárias 229 imagens.

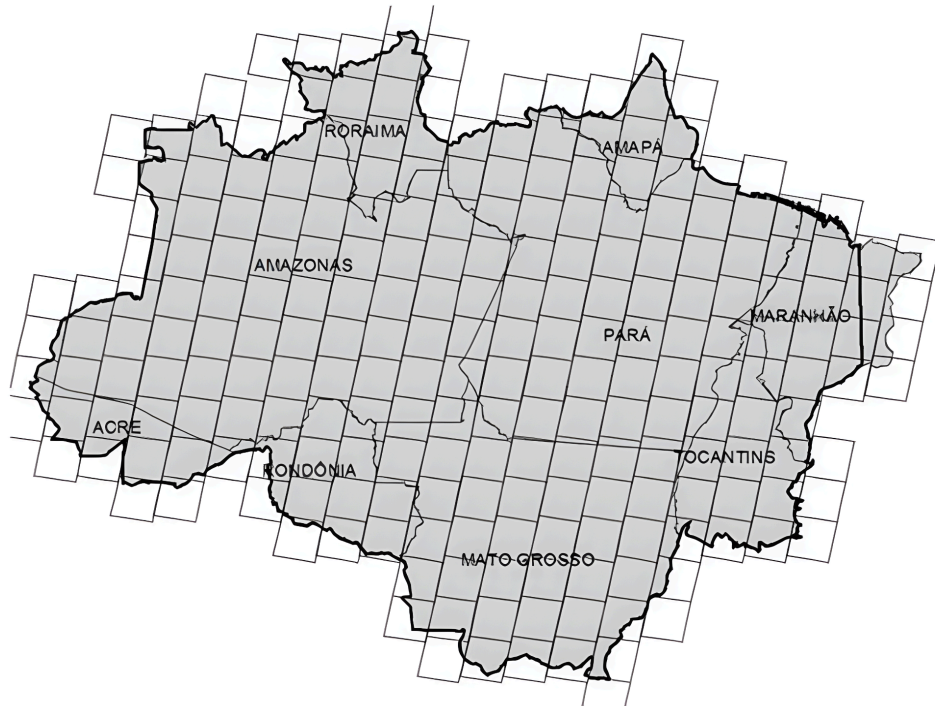


Figura 8: Localização das cenas Landsat que recobrem a Amazônia Legal Brasileira. Fonte: (DE ALMEIDA, CLAUDIO APARECIDO et al., 2022).

2.1.3.2 DETER

O DETER foi implementado em 2004 e tem um papel fundamental no apoio à fiscalização. Ele realiza o mapeamento da remoção e degradação florestal na Amazônia Legal Brasileira e em áreas de vegetação primária nas formações savânicas e florestais do bioma Cerrado. O DETER, diferente do PRODES, não é utilizado para o relatório anual e sim somente para avisos diários sobre alteração da vegetação (INPE, 2022).

Inicialmente, o sistema DETER utilizava o sensor MODIS, ele possuía uma resolução espacial de 250 m e identificava apenas desmatamentos acima de 25 hectares. Em 2015 o novo sistema chamado DETER-B entrou em operação e é empregado atualmente, passando a utilizar imagens dos satélites CBERS-4, 4A e Amazônia-1 do INPE, que possuem resolução espacial entre 56 e 64 metros. Isso possibilitou a categorização dos avisos em desmatamento e degradação e a detecção de áreas menores, de até 3 hectares. Além disso, esses satélites combinados oferecem uma taxa de revisita de 1 a 2 dias (INPE, 2023).

É importante ressaltar que o trabalho é realizado por uma equipe de especialistas que realiza diariamente a interpretação visual das imagens. O tempo de avisos mapeados pode não ocorrer de forma diária. Somando o tempo entre a passagem do satélite, mapeamento dos avisos, auditoria e inserção no banco de dados ocorre entre 48 e 72 horas, sem contar as variações dos dias trabalhados da equipe que realiza o trabalho (INPE, 2022).

2.1.3.3 TÉCNICAS DE CLASSIFICAÇÃO DO DESMATAMENTO

Tanto a classificação do desmatamento do PRODES e do DETER é realizada por meio de fotointerpretação visual de imagens, ou seja, são realizadas técnicas descritas a seguir para realce de aspectos de interesse e após isso os polígonos são delimitados manualmente por profissionais treinados. Para ambos os sistemas é utilizado o software desenvolvido pelo INPE chamado “TerraAmazon”. Abaixo estão descritas as duas principais técnicas especificadas no documento Metodologia Utilizada nos Projetos PRODES e DETER: (DE ALMEIDA, CLAUDIO APARECIDO et al., 2022).

Modelo Linear de Mistura Espectral (MLME): é uma técnica de processamento de imagens de satélite que permite dividir cada pixel das imagens em componentes distintos, como frações de solo, vegetação e sombra. Utilizando o MLME, é possível quantificar com precisão a proporção de cada um desses elementos em um único pixel. Isso é especialmente útil para identificar e analisar áreas de desmatamento e degradação ambiental. Ao diferenciar

a extensão do solo exposto, da cobertura vegetal e das áreas sombreadas, os analistas podem monitorar mudanças na paisagem, detectando eficientemente alterações causadas por atividades humanas.

Composição Colorida: Essa técnica envolve a combinação de diferentes bandas espectrais que são inseridas nos canais Vermelho, Verde, Azul (RGB) da imagem. Com isso, utilizando diferentes combinações é possível realçar aspectos de interesse facilitando a interpretação. Na metodologia PRODES são combinadas nos canais RGB as bandas infravermelho médio, infravermelho próximo e vermelho respectivamente. Já na metodologia DETER, são inseridas a banda do vermelho, infravermelho próximo e verde aos canais vermelho, verde e azul respectivamente.

A figura 9 exhibe a técnica de composição colorida utilizada no PRODES. Já a figura 10 exhibe a técnica de decomposição em frações (MLME) e a composição colorida empregada no DETER.

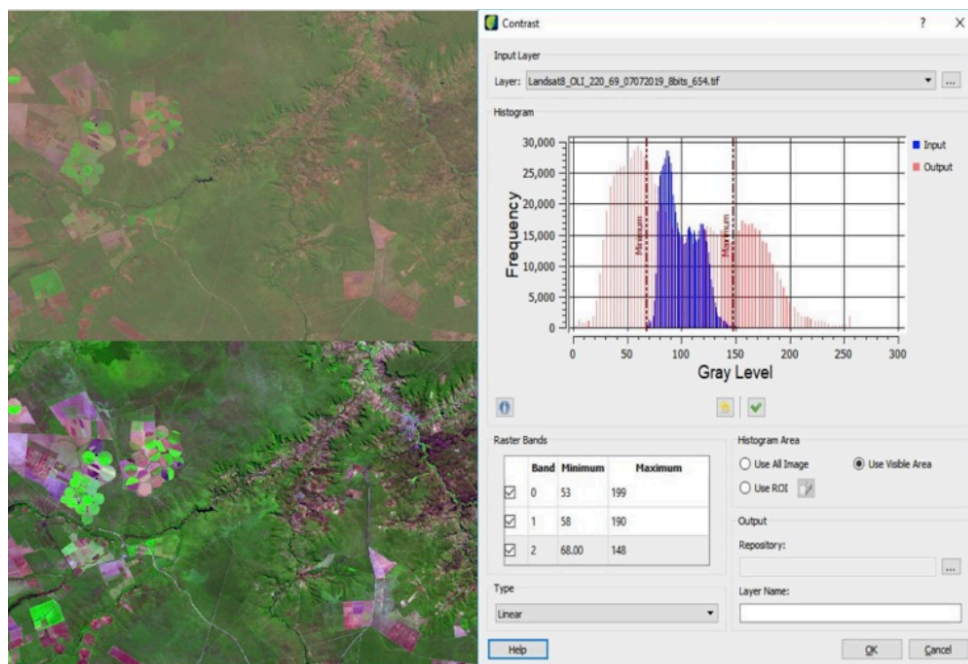


Figura 9: Aplicação da composição colorida PRODES. Fonte: (DE ALMEIDA, CLAUDIO APARECIDO et al., 2022).

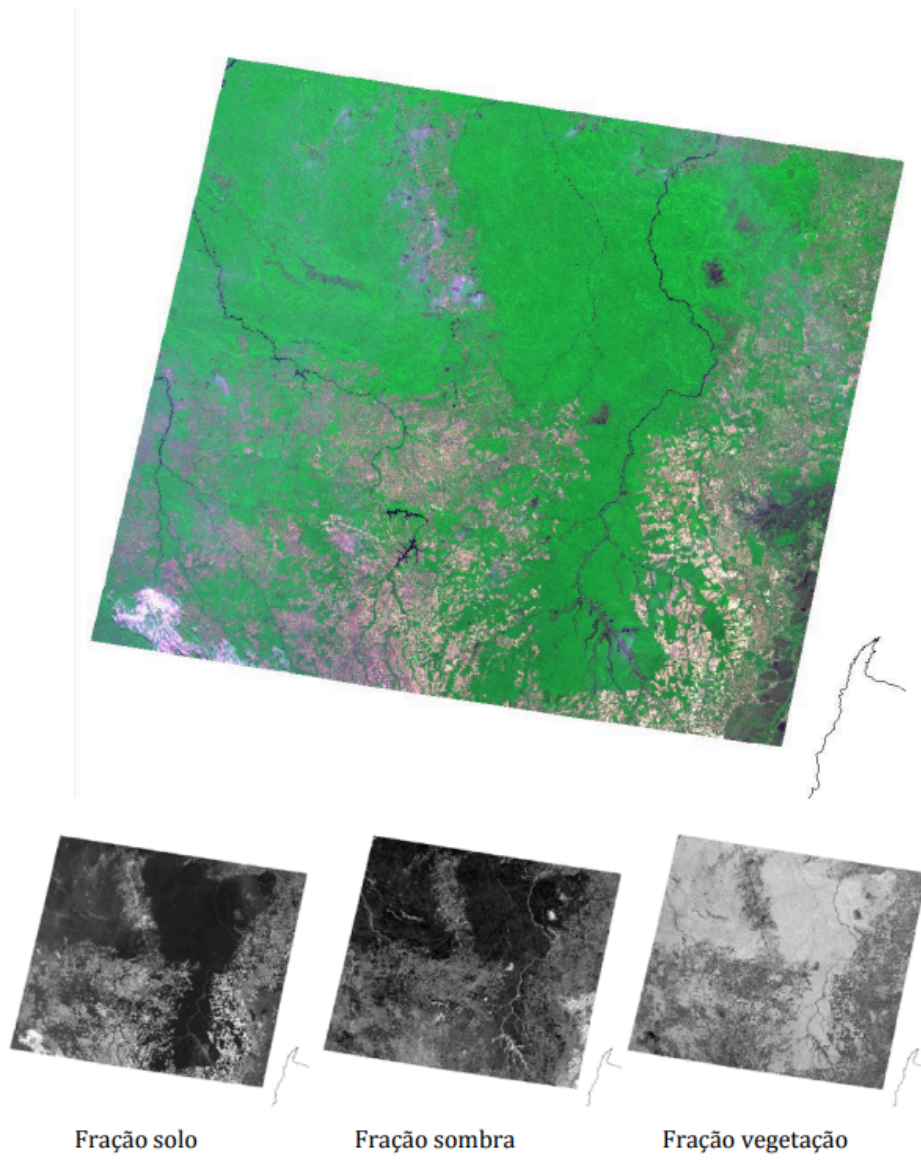


Figura 10: Aplicação da composição colorida e MLME DETER . Fonte: (DE ALMEIDA, CLAUDIO APARECIDO et al., 2022).

2.1.3.4 CLASSIFICAÇÃO DO DESMATAMENTO

Para uma classificação precisa do desmatamento, as seguintes categorias foram estabelecidas conforme relatório do Deter-B publicado por (DINIZ, CESAR GUERREIRO et al., 2015): 1) Desmatamento através de corte raso; 2) Desmatamento em áreas ainda com vegetação; 3) Áreas de mineração; 4) Áreas com degradação ambiental moderada; 5) Áreas com degradação ambiental intensa; 6) Áreas com marcas de queimadas; 7) Áreas de exploração madeireira seletiva regular; e 8) Áreas de exploração madeireira seletiva convencional. Na figura 11 são exibidas a imagem empregada para verificar a região de desmatamento, foi selecionado entre a melhor imagem de composição colorida ou MLME para melhor ilustração ao leitor.

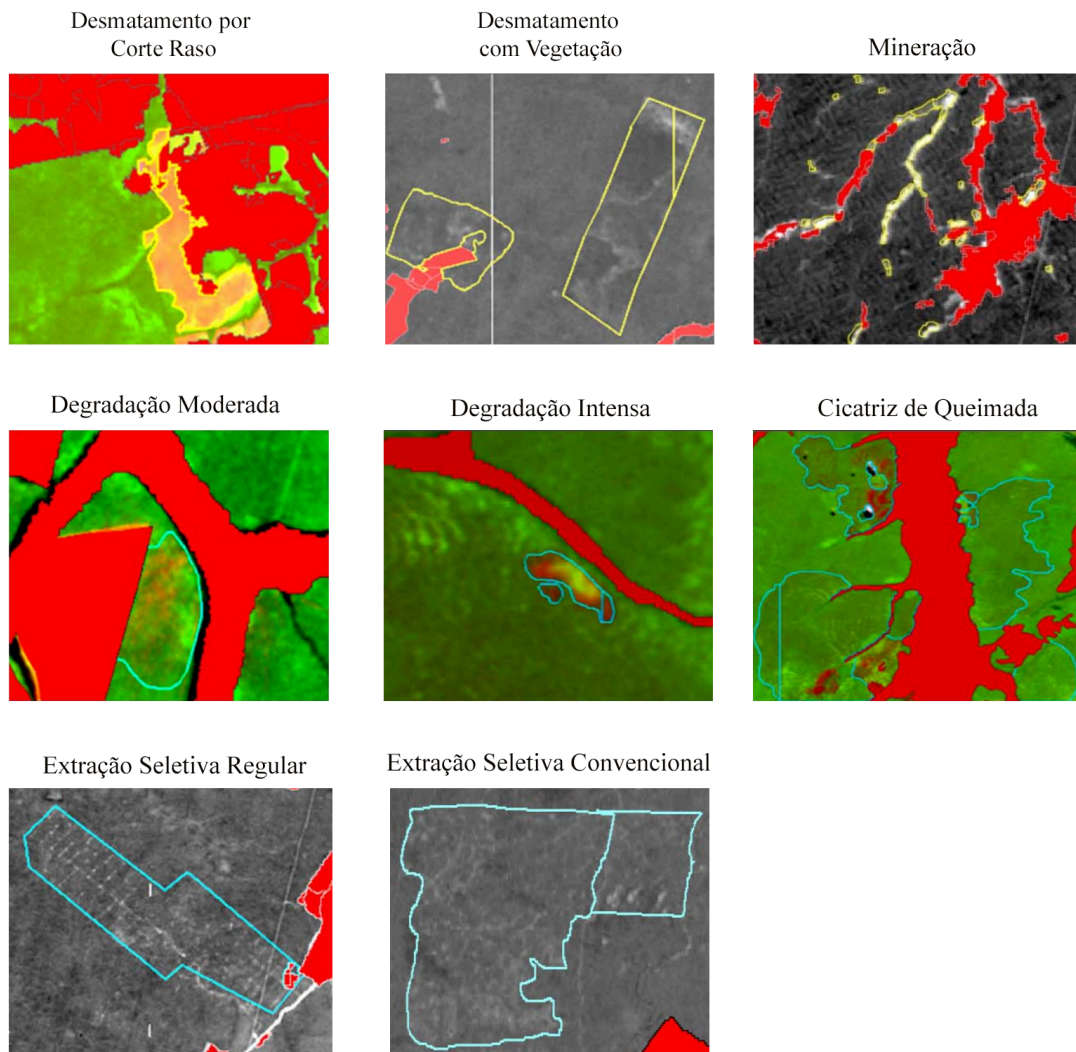


Figura 11: Categorias de Desmatamento. Adaptado. Fonte: (DINIZ, CESAR GUERREIRO et al., 2015).

2.1.4 SENSORIAMENTO REMOTO

2.1.4.1 SENTINEL 2

O programa Copernicus de observação da Terra foi lançado em 2014, iniciando o lançamento de satélites da série Sentinel. Seguindo essa iniciativa, em 2015 e 2017 os satélites Sentinel-2A e Sentinel-2B foram colocados em órbita respectivamente, operando em uma configuração sincronizada de 180° entre si. Cada satélite é equipado com um sensor multiespectral (MSI) que abrange 13 bandas espectrais, incluindo quatro bandas com resolução espacial de 10 metros (Azul, Verde, Vermelho e Infravermelho Próximo - NIR), seis bandas no espectro de borda vermelha e infravermelho de ondas curtas com resolução de 20 metros, e três bandas dedicadas à correção atmosférica com resolução de 60 metros. Além disso, o tempo de revisita são de 10 dias com somente um satélite e 5 dias com os dois satélites. (Agência Espacial Europeia, 2024).

2.1.4.2 AMAZONIA 1

O Amazonia 1 é o primeiro satélite de observação da Terra totalmente desenvolvido de forma autônoma no Brasil e tem como principal objetivo monitorar o meio ambiente na Amazônia e em outras áreas do país. O satélite faz parte da Missão Amazônia, projeto implementado pelo Instituto Nacional de Pesquisas Espaciais (INPE) e representa um avanço significativo da educação e desenvolvimento tecnológico espacial no Brasil, conforme destacado por Silva, Grande e Oliveira (2021), a sua produção de componentes tecnológicos brasileiros como o Subsistema de Propulsão, Caixa-Preta (responsável por armazenar imagens e áudios), Subsistema de Painéis Solares, Antena Transmissora de Dados e a *Câmera Wide Field Imager* (WFI), representaram entre 70% a 80% da fabricação do Satélite Amazonia 1, sendo um marco para a indústria aeroespacial brasileira. Ele foi lançado em órbita em 2021 e dentro de suas principais características possui um imageador óptico com 4 bandas espectrais (câmera com 3 bandas no espectro visível (Azul, Verde, Vermelho) e 1 banda no infravermelho próximo - NIR, além disso, possui uma taxa de revista de 5 dias e resolução espacial de 64 metros. (Ministério da Ciência, Tecnologia e Inovação, 2022).

2.1.5 INTELIGÊNCIA ARTIFICIAL

A inteligência artificial (IA) representa uma gama de algoritmos e técnicas de computação capazes de simular a inteligência humana, segundo Pacheco (2018), o teste de Turing serve para avaliar se o algoritmo pode ser considerado uma inteligência artificial ou não, para isso, ele deve ser capaz de enganar um ser humano ou fazer ele acreditar que ele é um humano. Este campo está presente em vários setores reformulando práticas convencionais e impulsionando inovações tecnológicas. No setor financeiro, por exemplo, a IA é uma ferramenta utilizada no combate a fraudes. No transporte, a IA aprimora sistemas de navegação e contribui para o desenvolvimento de veículos autônomos, trazendo uma visão das futuras dinâmicas de transporte. Na saúde, técnicas de segmentação semântica são utilizadas para detecção de doenças através de imagens médicas (MONTEIRO, ROSANGELA et al, 2022).

A inteligência artificial engloba diversas áreas, incluindo Aprendizado de Máquina (*Machine Learning*), Redes Neurais Artificiais (*Artificial Neural Networks*) e Aprendizado Profundo (*Deep Learning*). É importante destacar que o Aprendizado Profundo faz uso das Redes Neurais Artificiais, caracterizando-se pelo uso de múltiplas camadas ocultas, o que permite um aprendizado mais complexo e profundo (TAYE, 2023). Essas áreas podem ser visualizadas na figura 12.

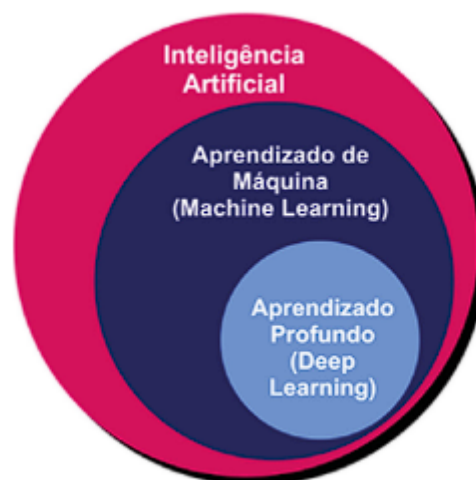


Figura 12: Áreas da Inteligência Artificial. Fonte: (MONTEIRO, ROSANGELA et al, 2022).

2.1.6 APRENDIZADO DE MÁQUINA

Aprendizado de máquina (*Machine Learning*) é um ramo da inteligência artificial que depende de grandes conjuntos de dados para aprender e fazer previsões sem ser diretamente programado para isso. Diferentemente dos métodos tradicionais, que necessitam de regras pré-definidas e dados de entrada para produzir uma saída, o aprendizado de máquina trabalha com os dados e as saídas observadas para descobrir as regras do processo durante o aprendizado. (PAIXÃO, SANTOS, et al, 2022). A figura 13 ilustra o modelo de programação clássica e aprendizado de máquina.

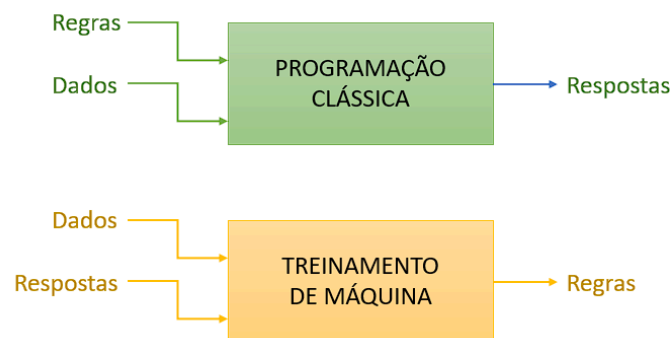


Figura 13: Programação Clássica x *Machine Learning*. Fonte: (GRANDO, 2022).

Os três tipos fundamentais de aprendizado de máquina são: Supervisionado, não supervisionado e aprendizado por reforço. O método supervisionado é aquele em que a rede recebe os dados de entradas rotulados para treinamento. Já o método não supervisionado, é aquele em que não são utilizados dados rotulados, ou seja, a rede durante o treinamento precisa encontrar padrões entre os dados de entrada. Por fim, o aprendizado por reforço, é caracterizado por um ambiente onde o algoritmo aprende por meio de tentativa e erro, utilizando recompensas ou penalidades para moldar a tomada de decisões. Existem diversos algoritmos que podem ser empregados dentro desses tipos de aprendizado como exemplos: modelos de regressão, algoritmos de classificação, árvores de decisão, métodos *bayesianos*, e redes neurais artificiais (JANIESCH, ZSCHECH, et al, 2021).

2.1.7 REDES NEURAIS ARTIFICIAIS

O desenvolvimento da tecnologia ao longo do tempo tem sido marcado por avanços significativos em várias áreas, sendo uma das mais notáveis o surgimento e a evolução das redes neurais artificiais. No início, as redes eram relativamente simples e suas capacidades limitadas. Porém, com o avanço da computação e o aumento do poder de processamento, essas redes tornaram-se cada vez mais complexas e capazes. Um aspecto notável dessa evolução é a forma como as redes neurais aprendem, inspiradas pelo funcionamento do cérebro humano, elas são capazes de aprender a partir de grandes quantidades de dados, identificando padrões e fazendo generalizações. (FLECK, LEANDRO et al, 2016).

Na figura 14 é possível visualizar o modelo mais simples de neurônio artificial. Ele é chamado de *Perceptron* e é caracterizado por um conjunto de 'n' entradas, cada uma dessas entradas é multiplicada por um peso específico, e os produtos resultantes são somados. Este somatório então passa por uma função de ativação que gera a saída resultante. (FLECK, LEANDRO et al, 2016).

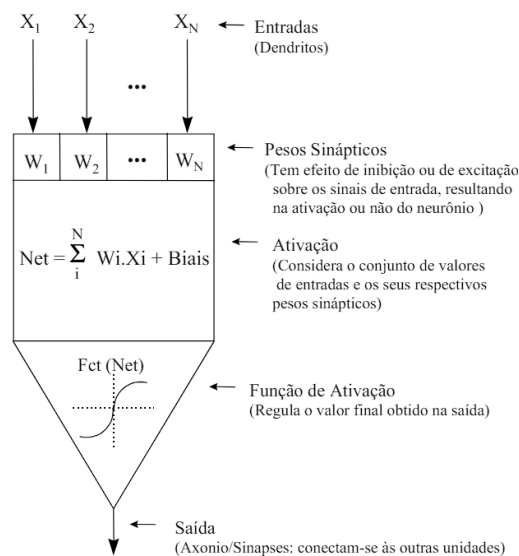


Figura 14: Modelo Perceptron. Fonte: (OSÓRIO, BITTENCOURT, 2000).

A partir do momento que é utilizado vários Perceptrons em uma rede, ela recebe o nome de *Multilayer Perceptron* (MLP), em muitas literaturas ela é também chamada de camada densa ou camadas totalmente conectadas. O MLP é constituído pelos elementos: camada de entrada, camada oculta ou intermediária e a camada de saída. A camada de entrada simplesmente repassa os dados para a próxima camada, a camada oculta, que pode ser uma ou várias dependendo da complexidade do modelo, em contextos de aprendizagem profunda, geralmente há múltiplas camadas ocultas. Nesta camada, uma função de ativação

não linear decide se e como as informações serão passadas adiante. Essa decisão é tomada com base nos sinais recebidos da camada anterior, que são ponderados pelos pesos que a rede adquiriu durante o treinamento. A camada de saída, por sua vez, recebe as informações das camadas intermediárias e gera a resposta final (DE MEDEIROS, TALLES HENRIQUE et al. 2005).

2.1.8 APRENDIZAGEM PROFUNDA

Quando uma rede neural artificial possui vários neurônios na camada oculta a rede é então considerada uma rede de aprendizado profundo (*Deep Learning*), por exemplo as redes neurais convolucionais, que possuem várias camadas de filtros para análise profunda e outros tipos de redes neurais que podem ser vistos na figura 15. Uma vantagem do aprendizado profundo é sua habilidade de extrair características importantes dos dados automaticamente, isso permite que esses tipos de modelos sejam aplicados de maneira eficaz a problemas complexos onde é difícil definir e extrair manualmente as características relevantes. Além disso, segundo Taye (2023), redes de aprendizado profundo foram muito bem sucedidas em tarefas de aprendizagem supervisionada, não supervisionada, híbrida e por reforço, corroborando a eficácia da utilização de redes neurais profundas. Entretanto, uma das desvantagens destes tipos de rede é que por sua estrutura ser complexa e possuir várias camadas, ela precisa de muito mais informações de entrada que outros modelos tradicionais de redes neurais, exigindo um poder computacional maior (TAYE, 2023).

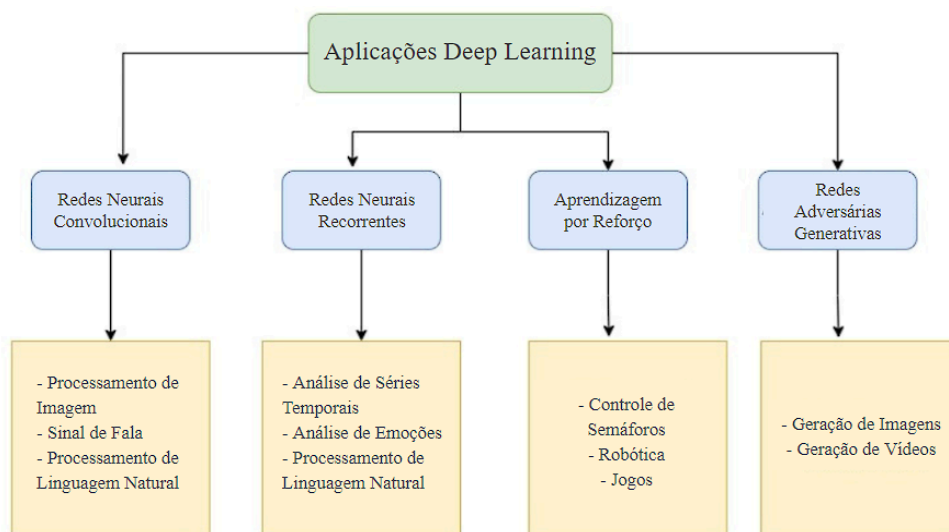


Figura 15: Áreas de Aplicação do *Deep Learning*. Adaptado. Fonte: (TAYE, 2023).

2.1.9 REDES NEURAS CONVOLUCIONAIS

Redes Neurais Convolucionais (*Convolutional Neural Networks - CNNs*) são um tipo de rede neural artificial especializada em processar dados com formato de matriz, como exemplo imagens. Elas são reconhecidas por sua capacidade de detectar automaticamente e aprender características relevantes sem supervisão humana direta. Essa capacidade as torna extremamente eficientes em tarefas como visão computacional (SOUZA, VICTOR et al, 2020).

2.1.9.1 CAMADA CONVOLUCIONAL

As CNNs são compostas pelas seguintes camadas: camada convolucional, camada de agrupamento ou redução e camada totalmente conectada. As camadas convolucionais utilizam filtros que extraem o mapa de características de uma imagem, através da operação de convolução é possível capturar os principais pontos, cada mapa corresponde a um canal dentro da matriz de imagem (SOUZA, VICTOR et al, 2020). Na figura 16 é exemplificado o cálculo da convolução, o resultado da convolução é o valor de saída [0] [0].

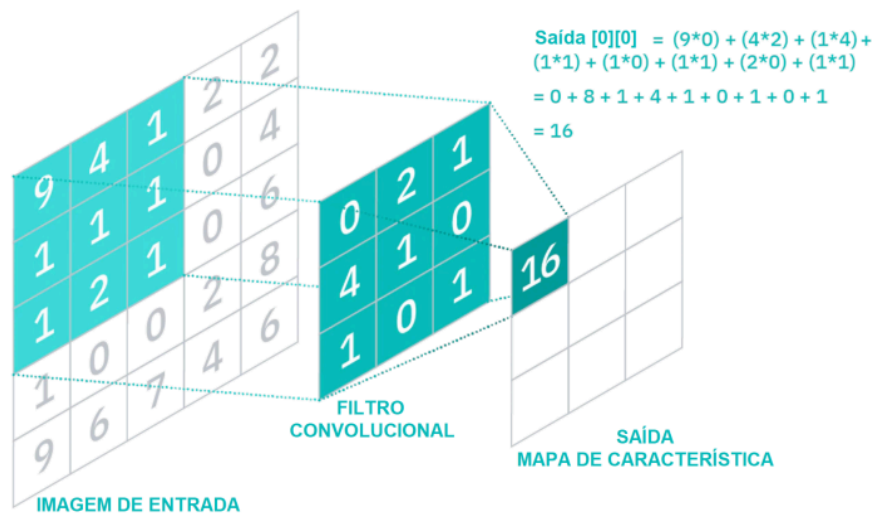


Figura 16: Exemplo de cálculo de Convolução. Fonte: (OLIVEIRA, 2022).

Na figura 17, demonstra-se o processo de convolução em uma rede neural convolucional. O filtro 3x3 utilizado nesse caso corresponde a cor azul escura dentro do bloco, ele percorre toda a imagem realizando operações de convolução e no final é obtido um mapa de características da imagem. Esse processo ocorre diversas vezes dentro da imagem com diferentes filtros que detectam diferentes detalhes da imagem como bordas, texturas ou padrões específicos. (SANTOS, 2021).

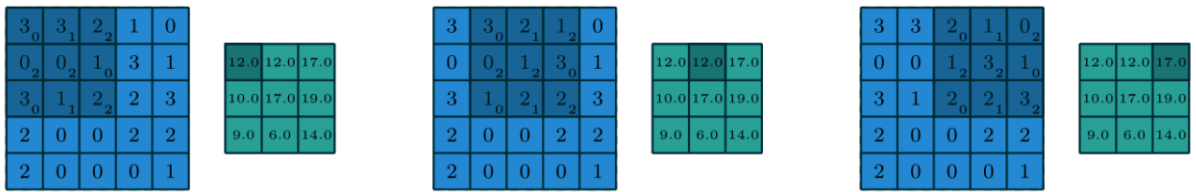


Figura 17: Exemplo de convolução na imagem. Fonte: (SANTOS, 2021).

Devido a natureza da convolução, a imagem geralmente perde a borda por conta da maneira como o filtro é aplicado sobre ela, diminuindo o tamanho da imagem. Para que o filtro seja aplicado de maneira uniforme, incluindo suas bordas, pode ser utilizado o parâmetro *padding* (Preenchimento) que adiciona pixels ao redor da imagem (geralmente com valor zero), permitindo que a dimensão de saída seja a mesma de entrada. O parâmetro *stride* (passo), é a forma com que filtro movimentar-se dentro da imagem, como exemplo na figura 18, foi aplicado um filtro 3x3, com *padding* 1 (uma borda de um pixel de valor 0 ao redor da imagem) e *stride* 1.

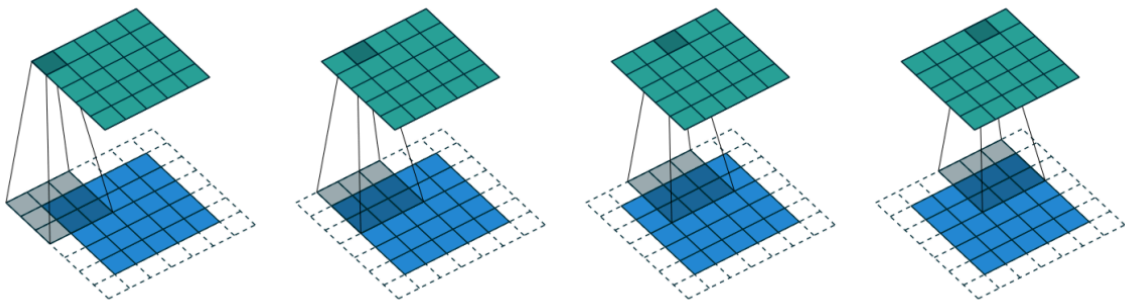


Figura 18: Operação de Convolução com *Padding*. Fonte: (SANTOS, 2021).

A quantidade de mapas de características em uma rede convolucional é igual ao número de filtros aplicados. Na figura 19 é exibido quatro mapas de características gerados através de diferentes filtros que são os quadrados em escala de cinza. É importante ressaltar que o símbolo “*” refere-se a operação de convolução e o símbolo “ σ ” corresponde a função de ativação. As principais funções de ativação empregadas após a aplicação do filtro são as ReLU (*Rectified Linear Units*), sigmóide e tangente hiperbólica. (PHIL, 2017).

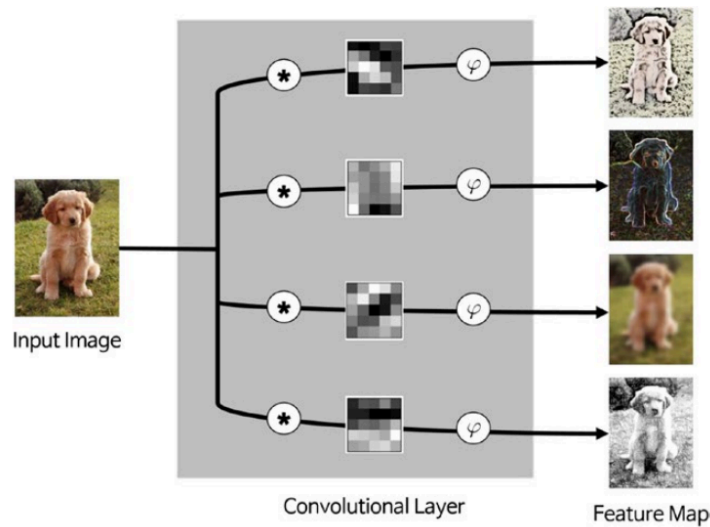


Figura 19: Mapas de características em uma convolução. Fonte: (PHIL, 2017).

2.1.9.2 CAMADA DE AGRUPAMENTO/REDUÇÃO

A camada de agrupamento ou redução, diminui o tamanho e largura da imagem e mantém o número de mapas de características. Na maioria das literaturas são aplicadas as funções de *Max-Pooling* (Agrupamento Máximo) e *Average-Pooling* (Agrupamento Médio). O *Max-Pooling* consiste em dividir o mapa em regiões (normalmente quadradas) e, em cada região, selecionar o valor máximo. Esse valor máximo é então mantido, enquanto os demais valores na região são descartados. O *Average-Pooling* utiliza a média dos valores. (SOUZA, VICTOR et al, 2020). No exemplo da figura 20 é realizada a operação de *Max-Pooling*:



Figura 20: Operação de *Max Pooling*. Fonte: (SANTOS, 2021).

2.1.9.3 CAMADA FINAL

A camada final, já estudada anteriormente, é a camada totalmente conectada, ou *Multilayer Perceptron* (MLP). A última camada totalmente conectada, que pode ser visualizada na figura 21, é responsável pela realização da classificação dos dados. Essa camada transforma as características extraídas pelas camadas convolucionais e de *pooling* em uma forma adequada para a tarefa de classificação.

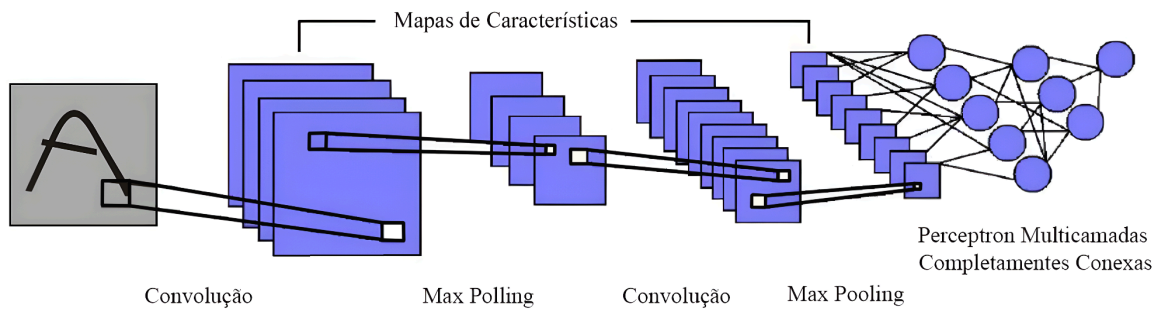


Figura 21: Rede Convolucional. Adaptado. Fonte: (FERREIRA, 2017).

2.1.9.4 CONVOLUÇÃO TRANSPONSTA

Como visto anteriormente, a natureza da convolução é realizar a redução do tamanho da imagem, aplicada principalmente na parte de *downsampling* (redução da resolução) na rede convolucional. Em alguns modelos de redes convolucionais, como exemplo a U-NET, após a operação de *downsampling* ocorre a operação de *upsampling* (aumento da resolução). Nesse processo, é empregado a conversão transposta para extração do mapa de características da imagem e aumento da resolução.

Na figura 22, o processo de convolução transposta é ilustrado detalhadamente. Durante este processo, o filtro desliza sobre cada pixel da imagem. Em cada posição, ocorre a multiplicação dos pesos do filtro pelo valor do pixel correspondente. Esta ação é repetida para cada pixel da imagem. A cada passo, o filtro gera uma matriz resultante da multiplicação do valor do pixel pelos pesos do filtro. Após o filtro ter percorrido toda a imagem, todas as matrizes resultantes são somadas para formar a saída final.

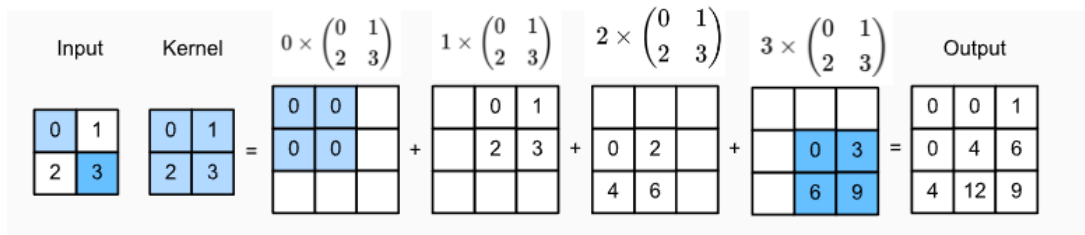


Figura 22: Cálculo de Convolução Transposta. Adaptado. Fonte: (ZHANG., LIPTON, et al, 2023).

Na convolução transposta, o *stride* refere-se à distância em que o filtro move-se para a geração da nova imagem ampliada. Por exemplo, na figura 23, um filtro de dimensões 3x3 é aplicado a uma imagem original de tamanho 2x2. Deste modo, obtém-se uma imagem ampliada com dimensões 5x5 pixels. É importante perceber que, nas imagens resultantes da figura 23, os valores da linha e coluna central devem ser somados devido à sobreposição dos filtros.

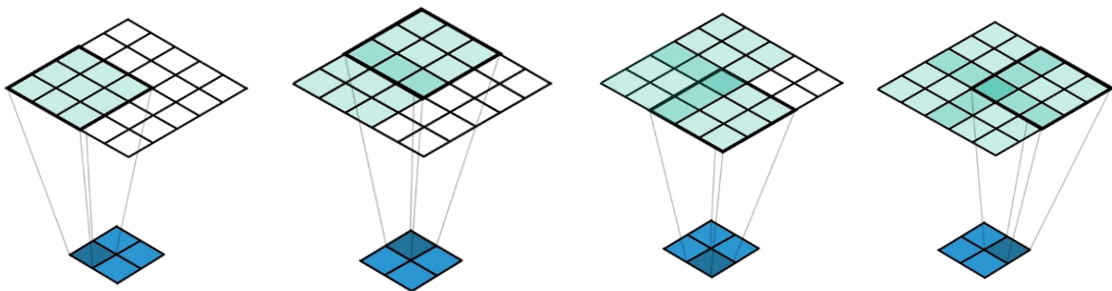


Figura 23: Convolução transposta em imagem de tamanho 2x2, filtro 3x3, *stride* 2. Fonte: (LANE, 2018).

2.1.9.5 CONVOLUÇÕES DILATADAS

Também chamadas de convoluções *atrous*, é uma técnica particularmente utilizada na segmentação semântica de imagens. Esta abordagem modifica a convolução tradicional ao inserir espaços no filtro de convolução, aumentando assim o campo receptivo sem um acréscimo no número de parâmetros ou na complexidade computacional. Sua principal vantagem é que através do parâmetro *rate* (avaliar) é possível ajustar o mapa de dilatação sem perder a resolução da imagem (Chen et al., 2017). Na figura 24, é possível visualizar como o filtro desliza pela imagem, já na figura 25, em (a) é exibido como uma estrutura de *encoder* ocorre sem e em (b) com a utilização da convolução dilatada.

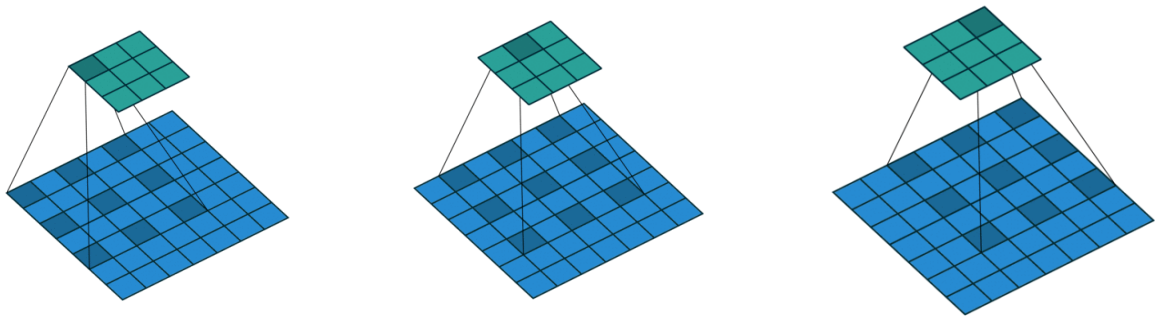
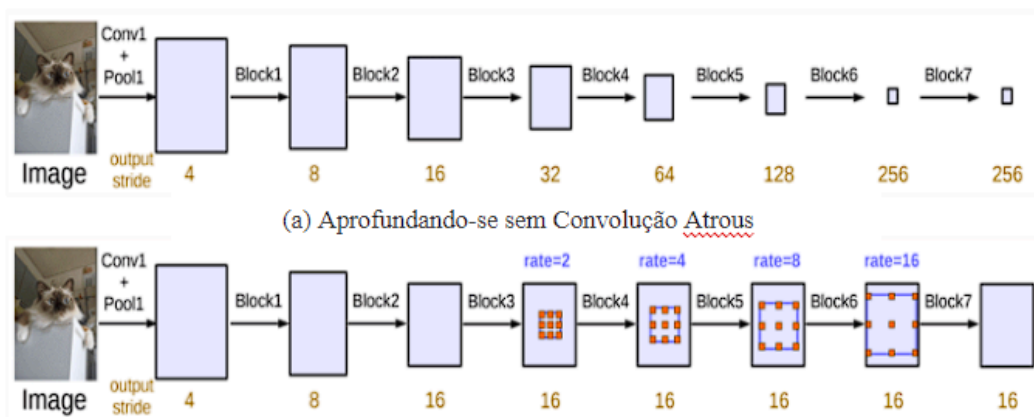


Figura 24: Exemplo de Convolução Dilatada. Fonte: (Chen et al., 2017).



(a) Aprofundando-se sem Convolução Atrous
 (b) Aprofundando-se com Convolução Atrous. Convolução Atrous com $rate > 1$ aplicado após o bloco 3 quando a $output_stride = 16$.

Figura 25: (a) Encoder (b) Convolução Atrous. Fonte: (Chen et al., 2017).

2.1.9.6 FUNÇÕES DE ATIVAÇÃO

As funções de ativação são uma parte importante das redes neurais convolucionais, elas determinam a saída de cada filtro convolucional aplicado. A operação de convolução é considerada uma operação linear, deste modo a utilização de uma função de ativação traz não linearidade ao processo tornando-o capaz de resolver soluções complexas, onde a relação entre variáveis ou características não é simples ou direta, e que não pode ser adequadamente descrita por modelos lineares ou fórmulas simples. (RODRIGUES, 2018).

As principais funções de ativação empregadas após a aplicação do filtro são as ReLU (*Rectified Linear Units*), sigmóide e tangente hiperbólica. (PHIL, 2017). Entretanto, conforme Leonardo Cardoso da Cunha (2020), a função ReLU pode ser treinada mais rapidamente trazendo melhores resultados e é ela a mais vista em redes neurais totalmente convolucionais.

Na figura 26, são exibidas as funções de ativações mencionadas.

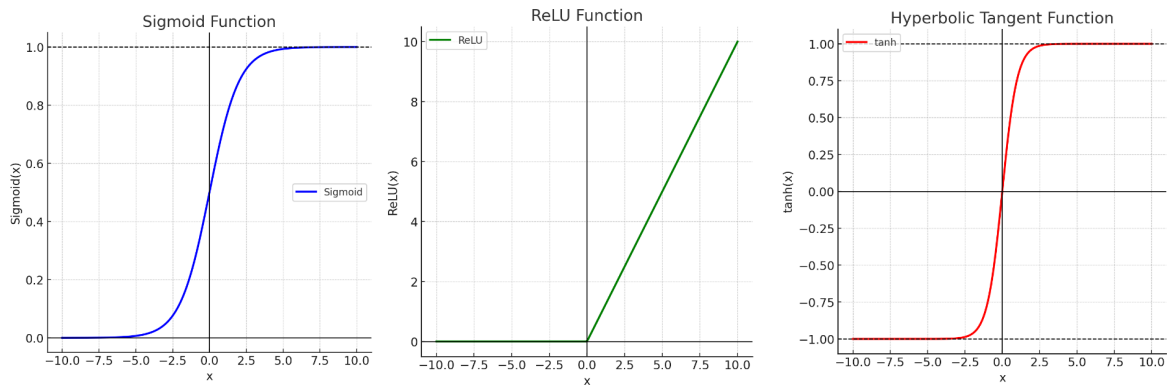


Figura 26: Funções de Ativação. Adaptado. Fonte (CUNHA, 2020).

A função de ativação mais empregada em camadas convolucionais, *ReLU*, é descrita como $f(x) = \max(0, x)$, onde ela zera todos os valores de ativação negativos e mantém os valores positivos na imagem (CUNHA, 2020). A figura 27 ilustra o seu funcionamento:

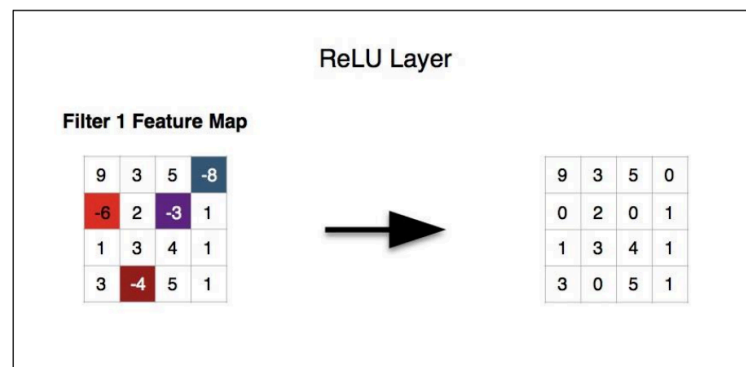


Figura 27: Funcionamento *ReLU*. Fonte (CUNHA, 2020).

2.1.10 SEGMENTAÇÃO SEMÂNTICA

A segmentação semântica, também chamada de *Fully Convolutional Neural Network* (Rede Neural Totalmente Convolutiva) refere-se à técnica de atribuir rótulos a cada pixel em uma imagem para indicar a classe à qual esse pixel pertence, proporcionando uma compreensão detalhada da distribuição espacial das diferentes classes na imagem. Em comparação, às redes neurais convolucionais tradicionais são geralmente utilizadas para realizar classificação de imagens inteiras, deste modo que a saída para uma imagem é uma única classe a que ela pertence (SILVA, 2021).

O que diferencia a segmentação semântica de outras redes neurais, é a substituição das

camadas totalmente conectadas por camadas convolucionais, por isso ela é denominada Totalmente Convolucional. As camadas convolucionais permitem que o modelo aceite imagens de entrada de tamanhos variados e gere mapas de probabilidade para cada pixel na saída, em vez de produzir uma única saída para a imagem inteira. (SILVA, 2021). Na figura 28, é possível visualizar uma representação desta atividade exibindo a imagem original ao lado esquerdo e sua versão segmentada ao lado direito.

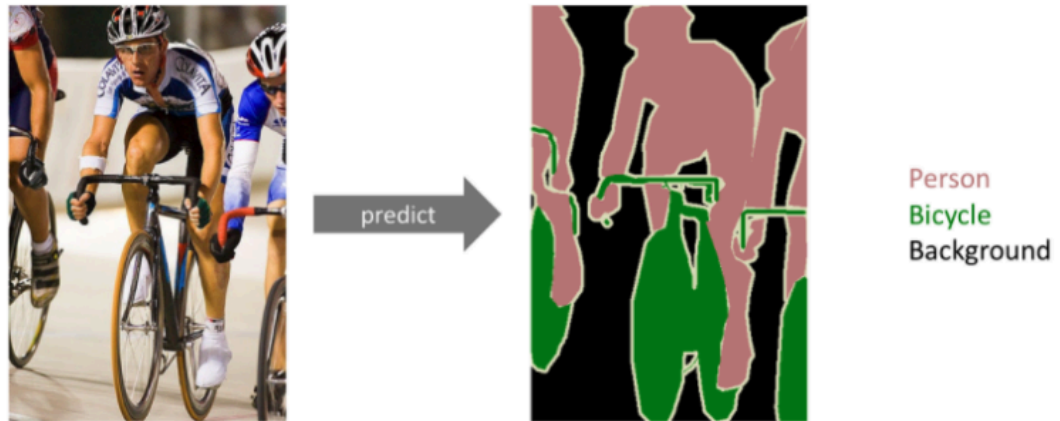


Figura 28: Segmentação Semântica. Fonte: (SILVA, 2021).

2.1.10.1 ARQUITETURA U-NET

A U-NET é uma arquitetura de rede neural convolucional que foi introduzida por Olaf Ronneberger, Philipp Fischer e Thomas Brox em 2015, focada inicialmente em tarefas de segmentação semântica em imagens biomédicas. Ela é caracterizada por dois principais componentes: *Encoder* (caminho da contração) e *Decoder* (caminho de expansão). O caminho de contração, utiliza camadas convolucionais para realizar o *downsampling* (redução da resolução), que é o processo de redução das dimensões da imagem. Esta etapa realiza a extração e compactação do mapa de características da imagem, facilitando a identificação de padrões relevantes em uma escala mais ampla. (RONNEBERGER, FISCHER, BROX, 2015). O nome “U-NET” deriva da forma como ocorre o processo descrito entre contração e expansão, na imagem a seguir fica evidente a arquitetura descrita na figura 29.

Durante a etapa de *downsampling* na imagem retirada do artigo “U-Net: Convolutional Networks for Biomedical Image Segmentation” de Ronneberger, Fischer e Brox, a imagem com dimensões $572 \times 572 \times 1$ foi inserida na rede e passou por convolução com filtro de tamanho 3×3 e função de ativação *ReLU*. Após a operação de convolução, ocorreu uma pequena redução na resolução da imagem, pois o filtro aplicado não pode se estender além

dos limites da imagem e não foi utilizado *padding* para manter as dimensões originais. Além disso, após a camada convolucional ocorre a operação *max pooling*.

Após a etapa de *downsampling*, é realizada a etapa de *upsampling* (Aumento da Resolução). Nesta parte do processo a rede convolucional realiza a expansão da imagem utilizando a técnica convolução transposta. Esta etapa é importante para reconstrução da imagem a partir das características compactas, permitindo a geração de uma segmentação precisa, com detalhes finos e bem definidos.

Outra operação de extrema importância que ocorre na arquitetura-NET é o copiar e concatenar, em inglês, chamado de *skip connections*. Essa função é usada para melhorar a precisão da segmentação ao reintroduzir informações detalhadas de resolução mais alta que seriam perdidas devido às operações de *pooling*.

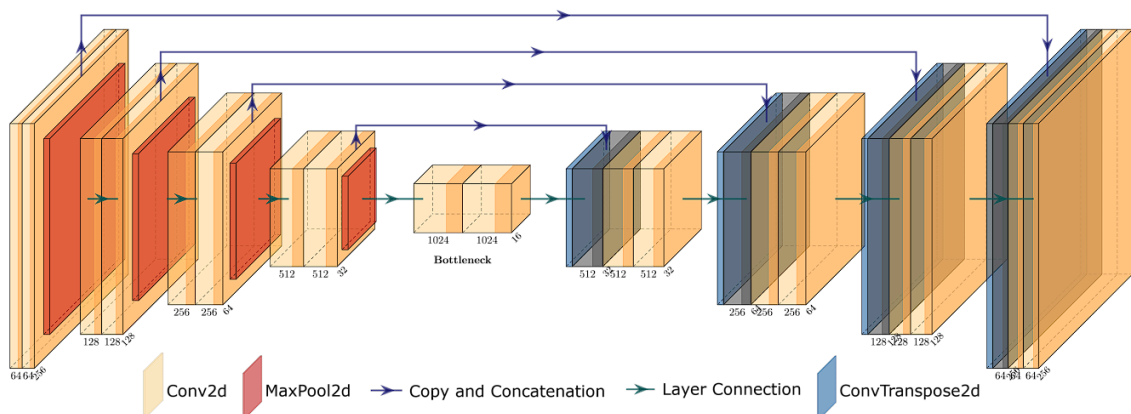


Figura 29: Arquitetura U-NET.. Fonte: (VILELA, 2024).

2.1.10.2 ARQUITETURA U-NET ++

A U-Net++ , introduzida por ZHOU, SIDDIQUEE, et al. em 2018, é uma arquitetura avançada de rede neural desenhada para aprimorar a segmentação de imagens. Esta rede é uma evolução da U-Net clássica, adicionando uma série de conexões densas e aninhadas entre as camadas, o que melhora a transferência de informações e gradiente através da rede. Além disso, a U-Net++ implementa um conceito de supervisão profunda, onde cada sub-rede no modelo é treinada para produzir um mapa de segmentação próprio, otimizando assim o processo de aprendizado. Através da figura 30, é ilustrado na cor preta a estrutura da U-NET padrão. As cores azul e verde representam as novas estruturas de *skip connections* e a cor

vermelha mostra a supervisão profunda descrita anteriormente.

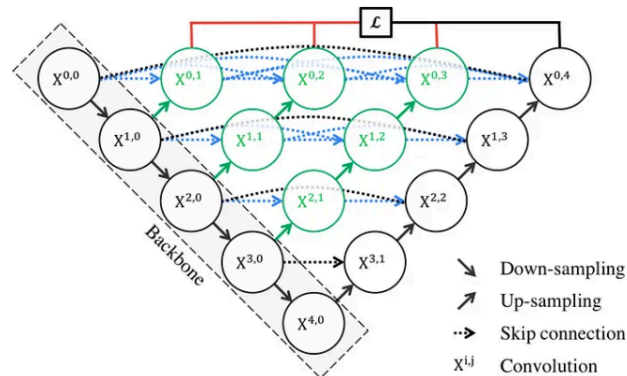


Figura 30: Arquitetura U-NET++. Fonte: (ZHOU, SIDDIQUEE, et al, 2018).

2.1.10.3 DEEPLAB V3 +

O DeepLab v3 +, desenvolvido por Chen et al. em 2018, é uma arquitetura de segmentação semântica avançada sucessora dos DeepLab v3, v2 e v1. Este modelo é notável por integrar um módulo *encoder-decoder* aprimorado ao já eficaz *Atrous Spatial Pyramid Pooling (ASPP)* do DeepLab v3, resultando em melhorias notáveis na precisão das bordas e detalhes finos da segmentação. Na figura 31, é possível ver o funcionamento do *Atrous Spatial Pyramid Pooling*, no capítulo 2.1.9.5 foi explicado o funcionamento da convolução dilatada ou convolução atrous, que adiciona espaços no filtro de convolução. Para o funcionamento da pirâmide, são basicamente várias convoluções dilatadas que ocorrem com diferentes taxas de *rate* para análise da imagem completa.

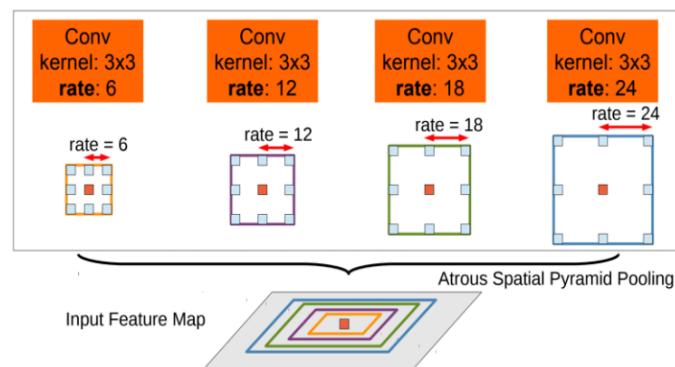


Figura 31: Funcionamento do *Atrous Spatial Pyramid Pooling*. Fonte: (Chen et al, et al, 2018).

A arquitetura do DeepLab v3+ é dividida em duas partes: A estrutura do *encoder*, que

utiliza o convoluções atrous para capturar informações em várias escalas da imagem, e o *decoder*, que refina os resultados da segmentação, concentrando-se em detalhes finos e bordas. A figura 32 ilustra o funcionamento de uma arquitetura DeepLabV3+.

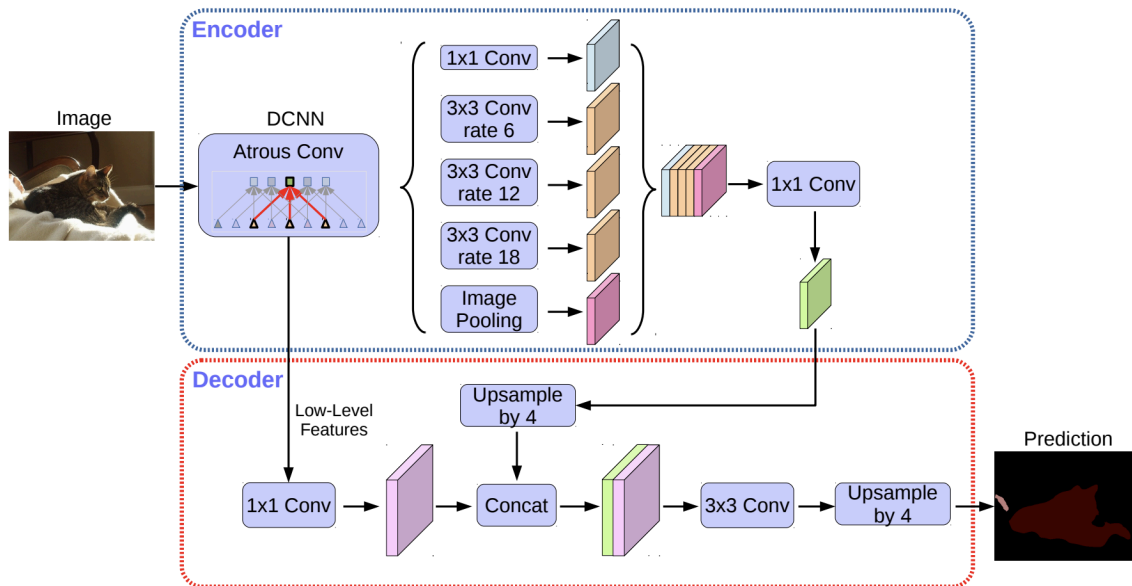


Figura 32: Arquitetura DeepLabV3+. Fonte: (Chen et al, et al, 2018).

3 METODOLOGIA

Nessa seção são detalhadas as etapas aplicadas para realizar a identificação das áreas de desmatamento da região da Amazônia através da segmentação semântica. O processo é dividido nas etapas de coleta de dados dos satélites Amazonia 1 e Sentinel-2, ajuste e preparação dos dados, desenvolvimento do modelo de segmentação semântica, avaliação de desempenho e descrição dos recursos de desenvolvimento empregados no trabalho. Todos os programas utilizados para desenvolvimento estão disponíveis no seguinte link do GitHub: https://github.com/eliassjr/Amazonia1_Sentinel2_Segmentation.git

3.1 COLETA DE DADOS

3.1.1 AQUISIÇÃO DE IMAGENS AMAZONIA 1

A aquisição de imagens da Amazônia foi realizada através do portal disponibilizado pela Divisão de Geração de Imagens do INPE. As imagens disponibilizadas pelo satélite possuem uma largura de faixa imageada de 850km, ou seja, cada imagem cobre uma enorme área de território. Para composição do dataset, foram selecionadas 10 imagens principais, e a partir dessas, realizaram-se recortes para gerar imagens menores de 256x256 pixels, o que corresponde a uma área de 16,38 km por imagem levando em conta que a resolução espacial do satélite é 64x64 metros por pixel. Esta abordagem permite uma análise focalizada nas regiões de interesse, além de respeitar as limitações de processamento do *hardware* utilizado. Na figura 33 é possível visualizar as imagens do satélite obtidas no site do INPE.

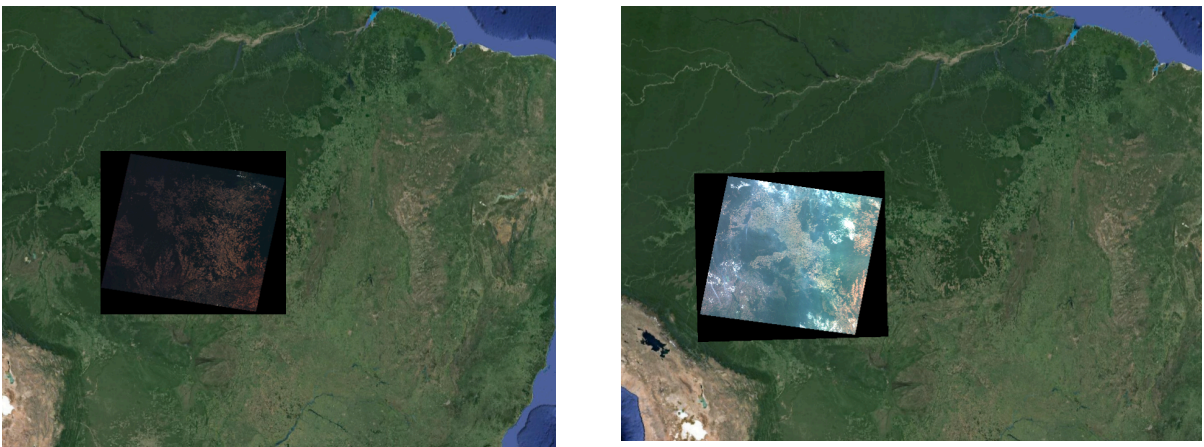


Figura 33: Imagens do satélite Amazonia 1. Fonte: INPE (2024).

As máscaras de desmatamento para treinamento da rede foram obtidas através da

plataforma Terra Brasilis, que é um site web desenvolvido pelo INPE para acesso a dados geográficos. A máscara é disponibilizada em formato shapefile, na figura 34 é possível visualizar a máscara do desmatamento total até o ano de 2023. Já na figura 35 é exibido a máscara em formato shapefile dentro do software utilizado para geoprocessamento QGIS.

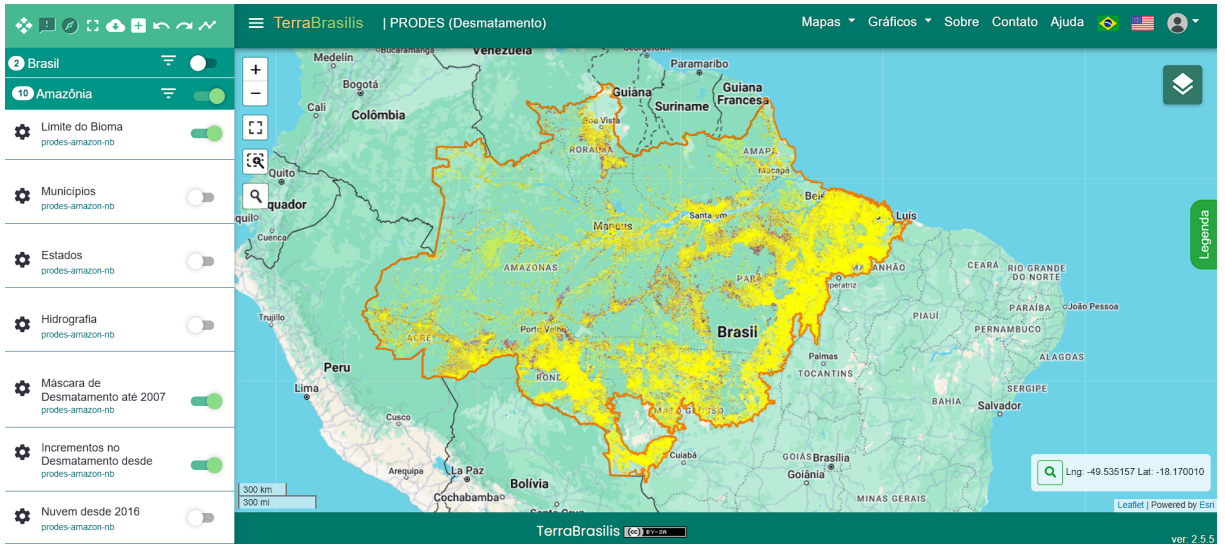


Figura 34: Plataforma TerraBrasilis. Fonte: Terra Brasilis (2024).

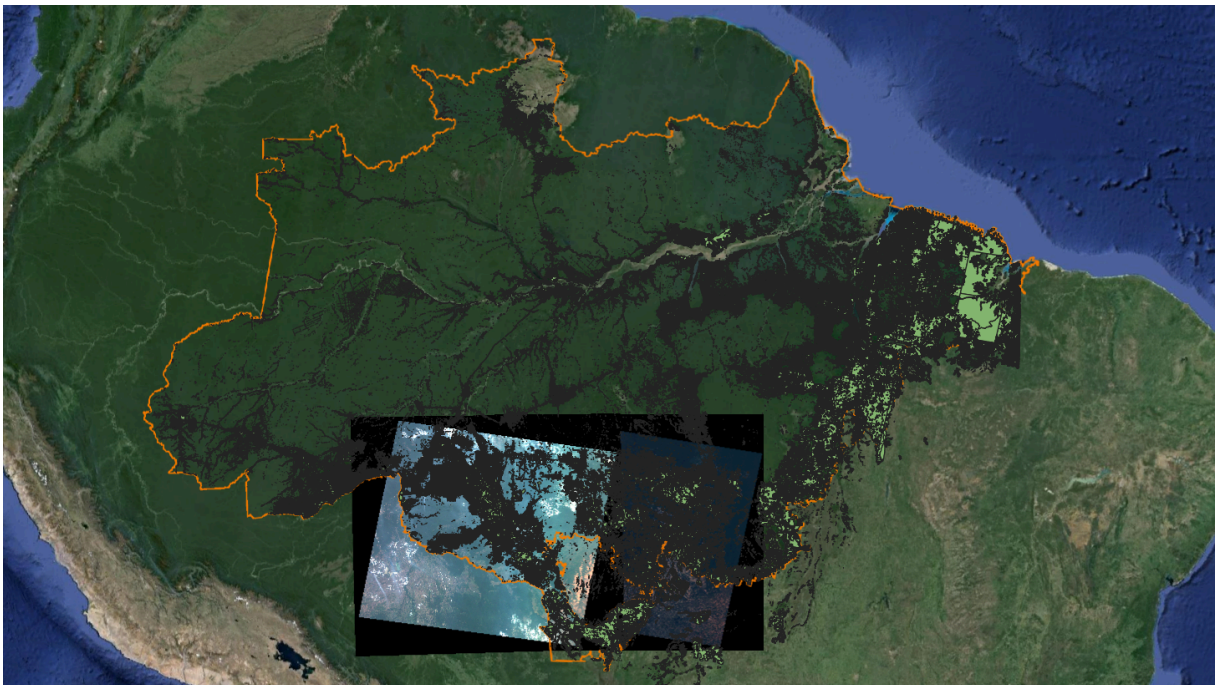


Figura 35: Imagens do satélite Amazonia 1 e shapefile de desmatamento. Fontes: Terra Brasilis (2024), INPE (2024).

O recorte das imagens foi realizado através de um programa criado em Python que utiliza a biblioteca Gdal para processamento de imagens georreferenciadas. Para a seleção das imagens, basta o usuário colocar a coordenada desejada que é realizado um recorte da imagem com dimensões 256x256 pixels, além disso, o programa realiza o recorte da mesma região do arquivo shapefile e o converte para uma imagem PNG, fornecendo assim tanto a imagem e sua máscara de desmatamento para o dataset. A figura 36 exibe o processo realizado, onde pode-se visualizar as imagens recortadas com suas máscaras de desmatamento em cor preta. Já na figura 37, é exibido uma região de desmatamento e sua respectiva máscara de desmatamento através do satélite Amazonia 1.

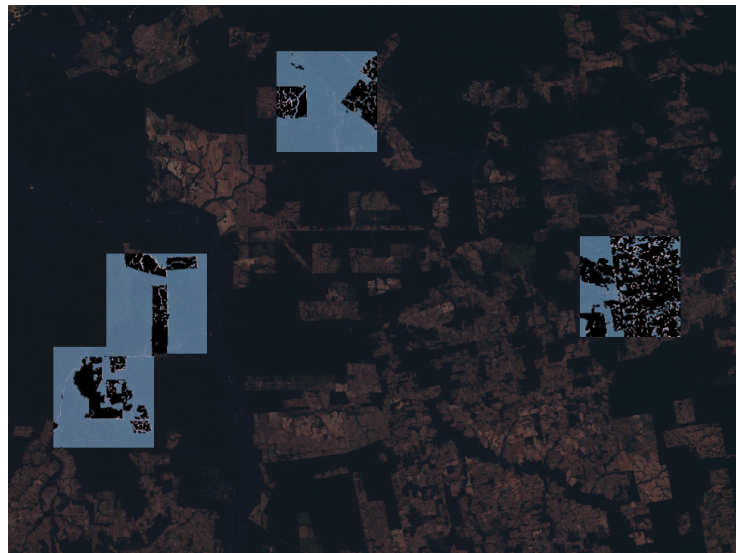


Figura 36: Imagens do satélite Amazonia 1 e suas máscaras de desmatamento. Fontes: Terra Brasilis (2024), INPE (2024).



Figura 37: Imagem do satélite Amazonia 1 e sua máscaras de desmatamento. Fonte: Terra Brasilis (2024), INPE (2024).

3.1.2 AQUISIÇÃO DE IMAGENS SENTINEL-2

Para as imagens do Sentinel-2, foi utilizado um dataset disponibilizado pelos autores José Mario Vicensi Grzybowski, Lucimara Bragagnolo e Roberto Valmir da Silva chamado *Amazon and Atlantic Forest image datasets for semantic segmentation*. Dentro desse arquivo possui imagens do Sentinel-2 com nível de imagens 2A, que é um dos níveis mais utilizados pois já passou por um processo de correção atmosférico. Os arquivos já estão convertidos em 8 bits conforme imagens utilizadas no Amazonia 1, além disso as imagens possuem dimensões 512 x 512 pixels, resolução espacial de 10 metros e as bandas disponibilizadas são vermelho, verde, azul e infravermelho próximo. O dataset possui no total 499 imagens e máscaras de treino, 20 de teste e 100 para validação. A figura 38 é um exemplo de imagem do dataset empregado no projeto.

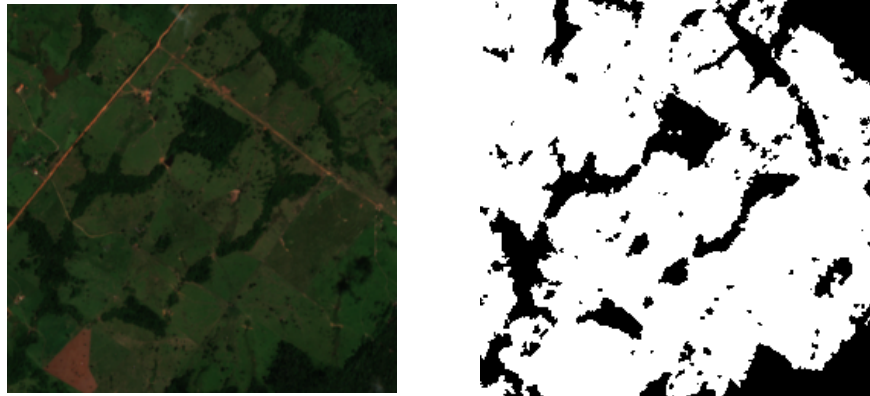


Figura 38: Imagem do satélite Sentinel-2. Fonte: (BRAGAGNOLO, LUCIMARA, et al, 2021)

3.2 AJUSTE E PREPARAÇÃO DOS DADOS

Para construção da rede neural, foi definido que as imagens de entrada devem possuir dimensões de 256x256 pixels, conter as bandas vermelho, verde e azul e formato 8 bits, deste modo, cada pixel pode corresponder a um valor de 0 a 255. Tanto as imagens provenientes do Amazonia 1 e Sentinel-2 foram submetidas a um processo de redimensionamento para atender às dimensões e bandas requeridas. Na seleção das imagens do Amazônia 1, critérios como baixa cobertura de nuvens, datas entre 2023 e 2024 e altas incidências de desmatamento foram considerados. Já dentro do dataset *Amazon and Atlantic Forest image datasets for semantic segmentation*, observou-se que muitas imagens possuem alta cobertura de nuvens, e além disso as máscaras estavam considerando as nuvens como uma área do desmatamento, deste modo foi necessária a seleção das melhores imagens. Após a escolha e realocação das melhores imagens, os datasets ficaram com 200 imagens para treino, 50 para teste e 50 para validação.

3.3 DESENVOLVIMENTO DO MODELO DE SEGMENTAÇÃO SEMÂNTICA

Para a construção da rede neural foi utilizado o modelo U-Net. A imagem de entrada possui dimensões de 256x256x3 (Altura x Largura x Canais). Durante o caminho de contração do modelo, a imagem é processada por operações convolucionais e de *pooling*, reduzindo sua dimensão para 16x16x1024 no ponto mais profundo da rede. No caminho de expansão, a imagem é progressivamente aumentada de volta à resolução original por meio de convoluções transpostas, alcançando 256x256 pixels com 64 canais. A saída é gerada após uma camada de convolução final que usa a função de ativação sigmóide, produzindo uma imagem de 256x256x1. Cada pixel dessa imagem final representa a probabilidade de desmatamento, onde 1 equivale a 100% de desmatamento e 0 a 0% de desmatamento. Foi definido que todos os valores iguais ou acima de 0,5 devem ser considerados como desmatamento e abaixo de 0,5 como não desmatamento.

3.4 AVALIAÇÃO DE DESEMPENHO

Para avaliar o desempenho do modelo de rede neural, utilizou-se as seguintes métricas: IoU (Intersecção sobre União), Acurácia (Accuracy), *Recall* (Revocação) e F1-Score. É importante compreender os diferentes tipos de predições que um modelo pode fazer, que incluem Verdadeiro Positivo (VP), Verdadeiro Negativo (VN), Falso Positivo (FP) e Falso Negativo (FN). Na figura 39, há uma ilustração prática dessas categorias, mostrando os casos corretamente classificados (VP e VN) e os incorretamente classificados (FP e FN).









	DOENTE 	SAUDÁVEL 
TESTE POSITIVO 	(VP) 	(FP) 
TESTE NEGATIVO 	(FN) 	(VN) 

Figura 39: Matriz de Confusão Exemplificada. Fonte: (SAMIR, 2024).

3.4.1 IOU

A métrica IoU, pode ser visualizada de forma intuitiva na figura 40, ela calcula a área

de intersecção (área comum) entre a previsão e o valor verdadeiro e divide isso pela área total da soma dessas duas. O valor de IoU varia de 0 a 1, onde 1 indica uma correspondência perfeita e 0 indica nenhuma sobreposição (MATOS, 2021). Em tarefas de segmentação semântica o IoU é mais utilizado, pois consegue identificar não só a localização mas também os contornos do objeto que tentou-se detectar, ele é diferente da acurácia que simplesmente conta o número de previsões corretas sobre o total, o que pode ser enganoso em imagens onde a maior parte do fundo é facilmente classificável, inflando o desempenho do modelo.

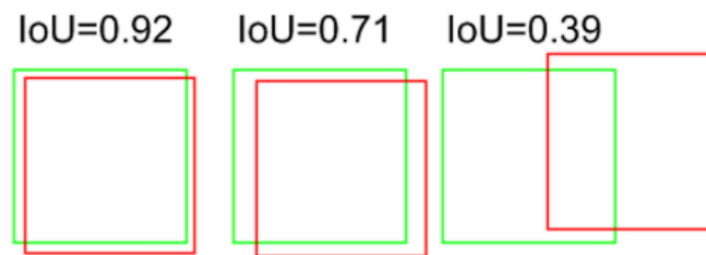


Figura 40: Intersecção sobre União. Fonte: (MATOS, 2024)

3.4.2 ACURÁCIA

A acurácia pode ser representada pela porcentagem de previsões corretas, tanto verdadeiros positivos quanto verdadeiros negativos, em relação ao total de previsões realizadas, este valor varia de 0 à 1 e quanto mais próximo do máximo melhor é a acurácia (CAMPOS, 2022). O uso individual dessa métrica para segmentação semântica é perigoso, pois ele pode inflar o desempenho em imagens que a maior parte da imagem é apenas o fundo que é simples detectar. A fórmula matemática está expressa abaixo:

$$\text{Acurácia} = \frac{\text{Verdadeiros Positivos} + \text{Verdadeiros Negativos}}{\text{Previsões Realizadas}}$$

3.4.3 RECALL

O recall é também conhecido como Revocação, é o número de verdadeiros positivos dividido pelo número de verdadeiro positivo mais falsos negativos. O Recall é importante em situações onde os falsos negativos são mais problemáticos do que os falsos positivos. Por exemplo, em testes médicos para doenças graves, é preferível ter alguns falsos positivos (e submeter indivíduos a exames adicionais) do que deixar de identificar pacientes que realmente têm a doença (CAMPOS, 2022).

$$\text{Recall} = \frac{\text{Verdadeiros Positivos}}{\text{Verdadeiros Positivos} + \text{Falsos Negativos}}$$

3.4.4 F1-SCORE

O F1 Score é uma métrica que combina Precisão e Recall através da média harmônica, ao invés de uma média aritmética simples, essa escolha é estratégica porque a média harmônica tende a penalizar desempenhos extremos nas componentes individuais. Ou seja, para alcançar um F1 Score alto, um modelo precisa apresentar tanto boa Precisão quanto bom Recall simultaneamente (CAMPOS, 2022). A fórmula matemática para calcular o F1 Score é a seguinte:

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

3.5 RECURSOS UTILIZADOS PARA DESENVOLVIMENTO

3.5.1 GOOGLE COLAB

Todo desenvolvimento dos códigos para aquisição de imagens, preparação, criação, treino e validação das redes neurais dos satélites Amazonia 1 e Sentinel 2 foram realizados através da plataforma Google Colab. Ela permite a programação e processamento em nuvem, disponibilizando as opções gratuitas de CPU e GPU limitada. Entretanto, através da aquisição do Google Colab Pro é possível utilizar 100 unidades de computação mensais. Os ambientes remotos utilizados foram o T4 GPU, com 16 GB de memória RAM na GPU, o consumo por hora foi de 1,76 unidades de computação e o A100 GPU, com 40 GB de memória RAM na GPU e consumo por hora de aproximadamente 16 unidades de computação.

3.5.2 TENSOR FLOW E KERAS

TensorFlow é uma das principais bibliotecas utilizadas para inteligência artificial, aprendizado de máquina e aprendizado profundo. Ele possui código aberto e desde seu lançamento pelo Google em 2015, rapidamente se tornou uma das ferramentas mais populares e influentes na área. Como uma biblioteca de computação numérica, o TensorFlow permite a construção de gráficos com dados complexos, facilitando cálculos intensivos e tornando possível operações de *machine learning* em uma variedade de dispositivos computacionais, como exemplos CPUs e GPUs (FALCÃO, MOREIRA, et al, 2019).

O TensorFlow 2 foi lançado em 2019, nessa versão ele vem junto com a biblioteca Keras, muito utilizado para criação de redes neurais profundas por conta de sua facilidade. No desenvolvimento do projeto, para criação da rede neural a biblioteca Keras foi fundamental principalmente para criação das camadas de convoluções, facilitando na definição de quantidade de filtros, tamanho, dentre outras características.

3.5.3 GDAL

O GDAL (*Geospatial Data Abstraction Library*), no português, Biblioteca de Abstração de Dados Espaciais, é utilizado no trabalho para recorte das imagens georreferenciadas e também para conversão dos shapefiles da forma vetorial para formato raster. Essa conversão é necessária para posterior criação das máscaras binárias de desmatamento ou não desmatamento.

3.5.4 QGIS

O software QGIS, exibido na figura 41, é uma abreviação para *Geographic Information System* (Sistema de Informações Geográficas). Ela mostrou-se uma plataforma muito eficiente e flexível para visualização e operações com dados georreferenciados. Todas as imagens principais utilizadas foram transportadas para dentro do software, além das máscaras de desmatamento obtidas na plataforma TerraBrasilis. Com o QGIS, foi possível realizar a conversão de sistemas de coordenadas e selecionar os melhores pontos de interesse de maneira simples para posteriormente inserir todas as coordenadas no programa desenvolvido para extração de imagens e máscaras de desmatamento. Além disso, todas as operações necessárias, como conversão de imagens, recorte e rasterização, são executáveis dentro do próprio software. Uma das funcionalidades é sua capacidade de fornecer o código Python correspondente às operações realizadas, o que facilita a programação no ambiente Google Colab.

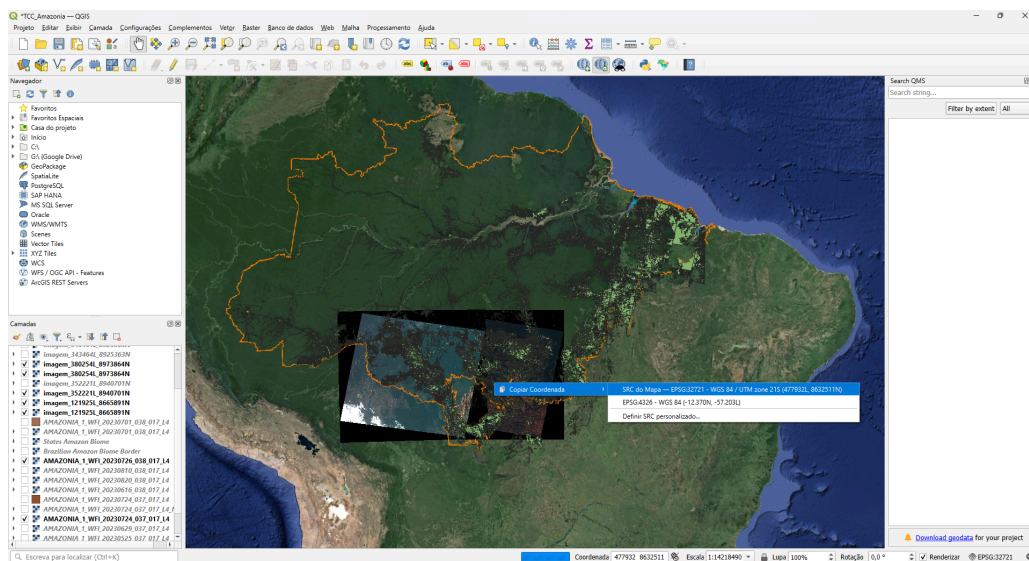


Figura 41: QGIS. Fonte: (AUTOR, 2024).

4 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Neste capítulo são evidenciados os testes e resultados obtidos através da segmentação semântica em imagens de satélite Amazonia 1 e Sentinel-2, com imagens de resolução espacial 64 e 10 metros respectivamente. O foco é avaliar a eficácia do modelo U-net para identificar áreas desmatadas, comparando entre as diferentes métricas de precisão IoU, Acurácia, F-Score e Recall.

4.1 PARÂMETROS UTILIZADOS

4.1.1 TAMANHO DO LOTE

O tamanho do lote (Batch Size) define o número de imagens que são propagadas pela rede de uma vez só, tamanhos de lotes maiores aceleram o tempo de processamento da rede neural, porém tamanhos de lotes menores podem levar a uma melhor generalização do modelo (GAO, ZHONG, 2020). Deste modo, para os testes com o dataset com 200 imagens utilizou-se 4 imagens por lote e nos testes onde foi utilizado a técnica de aumento de dados com 1200 imagens empregou-se 12 imagens por lote.

4.1.2 OTIMIZADORES

O otimizador dentro da rede neural tem o objetivo de ajustar os pesos para minimizar a função de perda *Dice loss*, que é a função aplicada no projeto, ela basicamente guia o processo de aprendizado. Os dois otimizadores aplicados para testes foram o Adam e sua versão melhorada Nadam, que pode levar o modelo a uma convergência mais rápida e a um treinamento mais estável.

4.1.3 INICIALIZADORES

O inicializador padrão de pesos nas camadas convolucionais do Keras é o Glorot, também chamado Xavier e leva o nome do seu criador, ele é projetado para equilibrar a variação dos gradientes ao inicializar os pesos, calculando-os com base na média do número de entradas e saídas das camadas, o que ajuda a evitar a explosão ou o desaparecimento dos gradientes durante o treinamento (GLOROT, BENGIO, 2010). Já o inicializador He, desenvolvido por HE et al. (2015), é otimizado para redes que utilizam a função de ativação ReLU, sendo efetivo no projeto desenvolvido, realizando o ajuste dos pesos com base apenas no número de entradas, para prevenir o problema dos gradientes que desaparecem, que são típicos desta função de ativação. Ambos inicializadores, Glorot e He foram empregados nos

testes realizados.

4.1.4 TAXA DE APRENDIZADO

A taxa de aprendizado, ou *learning rate*, foi definida inicialmente como 0,01, mas ela pode chegar até um valor mínimo de 0,0001. Para isso, foi utilizado o recurso `ReduceLROnPlateau`, que monitora o *validation loss* e realiza a redução da taxa de aprendizado em um fator de 0,1 após 3 épocas caso o modelo não apresentar melhoria.

4.1.5 ÉPOCAS E PARADA ANTECIPADA

Para treinamento de modelos com 200 imagens utilizou-se o limite de 64 épocas máximas de treinamento, e para 1200 imagens 128 épocas. Esse valor foi determinado para não gastar recursos de computação desnecessários, haja visto que na maioria dos modelos que chegaram perto desses valores a redução do *validation loss* ocorreu de forma ínfima. Além disso, outro recurso de supervisão empregado para evitar gastos computacionais desnecessários foi o parada antecipada com valor de 8 épocas, caso o modelo não apresenta melhora após essa quantidade de épocas o treinamento é encerrado.

4.1.6 AUMENTO DE DADOS

A técnica de aumento de dados (*Data Augmentation*) foi aplicada em alguns testes dos modelos. Isso foi necessário devido a limitação de imagens no banco de dados. Foram cinco operações realizadas para cada imagem: Giro horizontal, giro vertical, distorção óptica, distorção de grid e transformação elástica. Deste modo, o banco de imagens de treino subiu de 200 para 1200.

4.2 ANÁLISE DO DESEMPENHO

Todos os programas e conjuntos de dados empregados para treinamento, testes e validação estão disponíveis no seguinte link do GitHub: https://github.com/eliassjr/Amazonia1_Sentinel2_Segmentation.git. Realizou-se um total de 16 testes, sendo 8 com imagens do satélite Amazonia 1 e 8 com imagens do satélite Sentinel-2. Os resultados desses testes estão detalhados nas Tabelas 1 e 2 para o Sentinel-2, e nas Tabelas 3 e 4 para o Amazonia 1.

A análise dos resultados dos testes para os satélites mostram detalhes importantes sobre a eficácia resultante devido a variação dos parâmetros de treino entre diferentes

inicializadores de pesos e otimizadores, bem como o impacto do *data augmentation* e do volume de imagens de treino. O Sentinel-2, com uma resolução de 10 metros por pixel, fornece dados de alta qualidade que corroboram para uma segmentação precisa. Observou-se em seus resultados nos testes 5 a 8, que utilizaram *data augmentation* e um aumento número de imagens de treino (de 200 para 1200), uma pequena melhora nas métricas de IoU, Acurácia e F-Score comparados aos testes 1 a 4 sem *data augmentation*. Além disso, evidenciou-se que a variação dos inicializadores de pesos e otimizadores para os modelos não demonstraram eficácia evidente na melhora dos resultados.

Sentinel-2 (10MP)								
Teste	Data Augmentation	Imagens de Treino	Inicializadores de Pesos	Otimizadores	IoU	Acurácia	F-Score	Recall
1	Não	200	Glorot	Adam	0.78857	0.91587	0.87778	0.97470
2	Não	200	Glorot	Nadam	0.84349	0.94216	0.90475	0.91195
3	Não	200	He	Adam	0.88352	0.96281	0.92641	0.96526
4	Não	200	He	Nadam	0.86981	0.95192	0.92765	0.91938
5	Sim	1.200	Glorot	Adam	0.90380	0.96901	0.94720	0.95780
6	Sim	1.200	Glorot	Nadam	0.90748	0.97088	0.94730	0.95598
7	Sim	1.200	He	Adam	0.89639	0.96430	0.94434	0.95571
8	Sim	1.200	He	Nadam	0.90349	0.96662	0.94947	0.96250

Tabela 1: Resultados Sentinel-2. Fonte: (AUTOR, 2024).

Na tabela 2, observa-se o tempo de processamento, última época de treino antes da parada antecipada e a GPU empregada para testes. A utilização da GPU Tesla A100, especialmente nos testes que envolviam um maior volume de imagens, reduziu o tempo de processamento, permitindo um treinamento mais eficiente mesmo com o aumento da carga computacional.

Sentinel-2 (10MP)							
Teste	Tempo de Processamento	Última Época	GPU	Teste	Tempo de Processamento	Última Época	GPU
1	10 min	18	T4	5	40 min	27	T4
2	13 min	16	T4	6	51 min	125	A100
3	28 min	51	T4	7	14 min	31	A100
4	28 min	46	T4	8	20 min	48	A100

Tabela 2: Resultados Sentinel-2. Informações Computacionais. Fonte: (AUTOR, 2024).

O teste número 6 obteve maiores valores de IoU e acurácia, ele teve no total 125 épocas de treinamento e foi finalizado devido ao *Early Stopping*. Entretanto, é importante observar que a partir de um momento os valores de IoU e acurácia evoluem de forma ínfima com o tempo, trazendo uma utilização desnecessária das unidades de computação disponíveis. Em comparação, os modelos 5, 7 e 8 trouxeram resultados parecidos com um menor gasto computacional. A partir da figura 42, é possível observar a evolução do desempenho do modelo de rede neural número 6 ao longo das épocas. Adicionalmente, a figura 43 ilustra os resultados da segmentação realizados com este modelo.

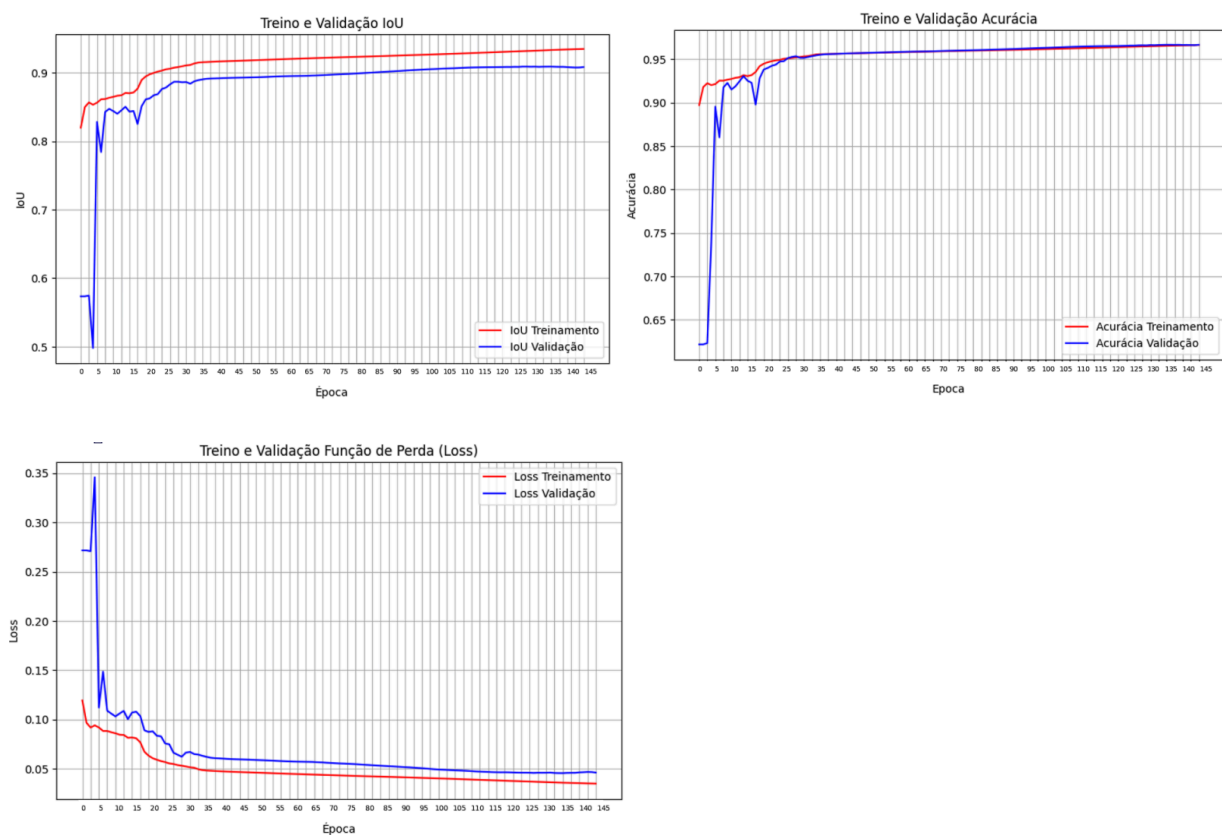


Figura 42: Treino e Validação IoU, Acurácia e Loss. Sentinel-2 Fonte: (AUTOR, 2024).

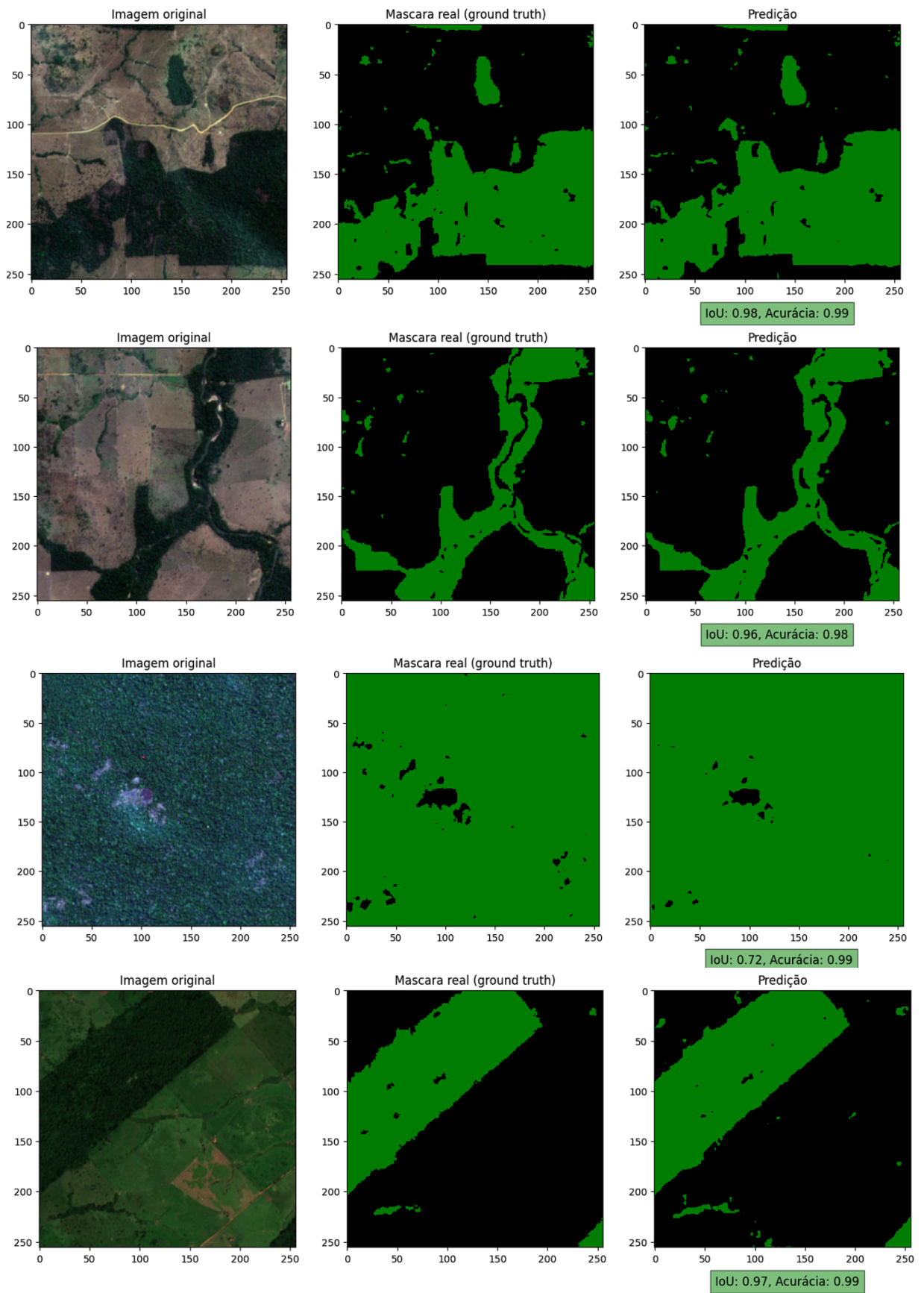


Figura 43: Segmentação Semântica. Sentinel-2 Fonte: (AUTOR, 2024).

O satélite Amazonia 1, tem uma resolução de 64 metros por pixel, o que implica em dificuldades adicionais devido à menor precisão das imagens. Através dos resultados exibidos nas tabelas 3 e 4, a análise revela que as métricas como IoU e Acurácia são consistentemente inferiores às obtidas com o Sentinel-2. Por exemplo, o melhor IoU para o Amazonia 1 é de 0.81427, enquanto o Sentinel-2 alcançou 0.90748. Além disso, similar ao Sentinel-2, o aumento nas imagens de treino melhoraram um pouco o desempenho dos modelos também para o Amazonia 1, como exemplo, a acurácia melhorou de 0.89911 para 0.91194 com o data augmentation.

Amazonia 1 (64MP)								
Teste	Data Augmentation	Imagens de Treino	Inicializadores de Pesos	Otimizadores	IoU	Acurácia	F-Score	Recall
1	Não	200	Glorot	Adam	0.77635	0.88860	0.85706	0.91139
2	Não	200	Glorot	Nadam	0.79302	0.89937	0.86366	0.90556
3	Não	200	He	Adam	0.79158	0.89911	0.86307	0.89709
4	Não	200	He	Nadam	0.79999	0.90399	0.86698	0.90864
5	Sim	1.200	Glorot	Adam	0.81443	0.91177	0.87572	0.91572
6	Sim	1.200	Glorot	Nadam	0.80999	0.90880	0.87325	0.91562
7	Sim	1.200	He	Adam	0.81093	0.91018	0.87353	0.91108
8	Sim	1.200	He	Nadam	0.81427	0.91194	0.87559	0.91547

Tabela 3: Resultados Amazonia 1. Fonte: (AUTOR, 2024).

Amazonia 1 (64MP)			
Teste	Tempo de Processamento	Última Época	GPU
1	21 min	35	T4
2	39 min	64	T4
3	16 min	18	T4
4	20 min	29	T4
5	43 min	121	A100
6	41 min	86	A100
7	14 min	29	A100
8	62 min	69	L4

Tabela 4: Resultados Amazonia 1. Informações Computacionais. Fonte: (AUTOR, 2024).

A comparação dos testes realizados revela que os modelos com *Data Augmentation* apresentaram desempenho superior. No entanto, dentro desses modelos que utilizaram essa técnica, as métricas são bastante equilibradas, tornando difícil a escolha do melhor modelo. Para efeitos de análise, a imagem a seguir mostra o desempenho do modelo da rede número 5 ao longo das épocas. Similarmente ao observado no treinamento com o Sentinel-2, verifica-se que após determinada época, os ganhos adicionais em termos de treinamento da rede neural se tornam pequenos, resultando em um consumo de recursos computacionais que pode ser considerado desnecessário. Os resultados da segmentação semântica através do satélite Amazônia 1 pode ser visualizado nas imagens da figura 45.

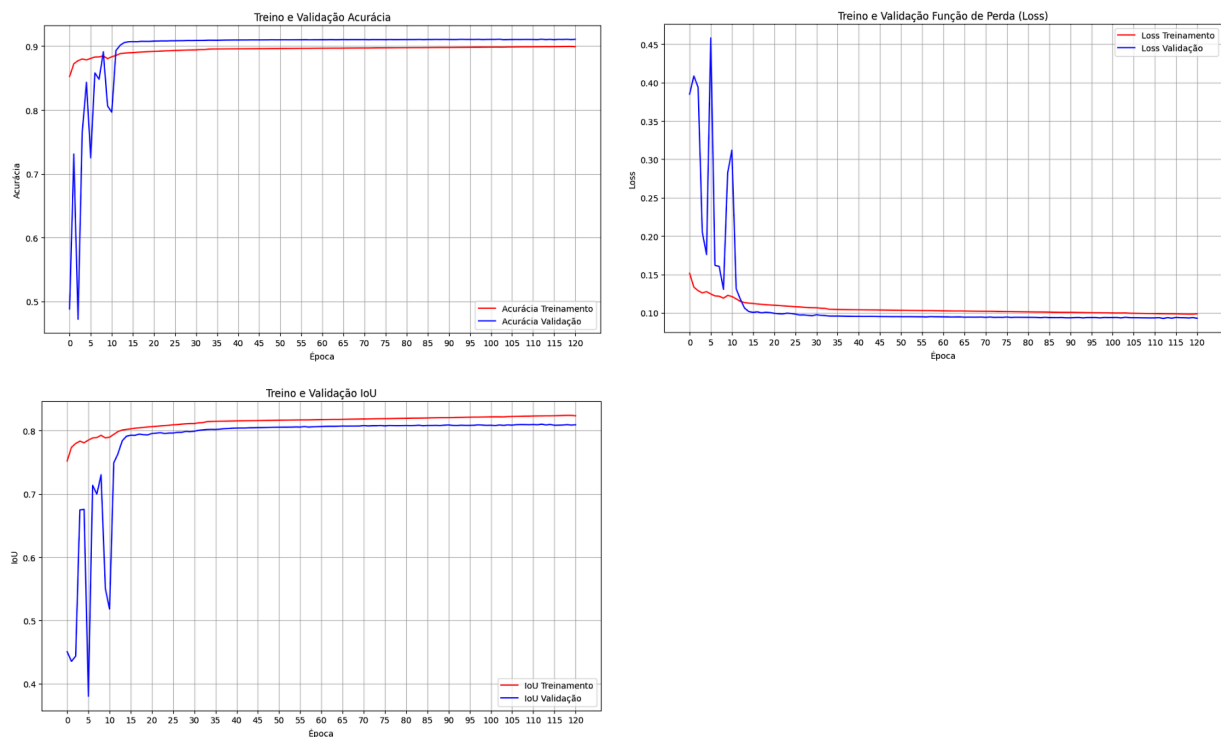


Figura 44: Treino e Validação IoU, Acurácia e Loss. Amazonia 1. Fonte: (AUTOR, 2024).

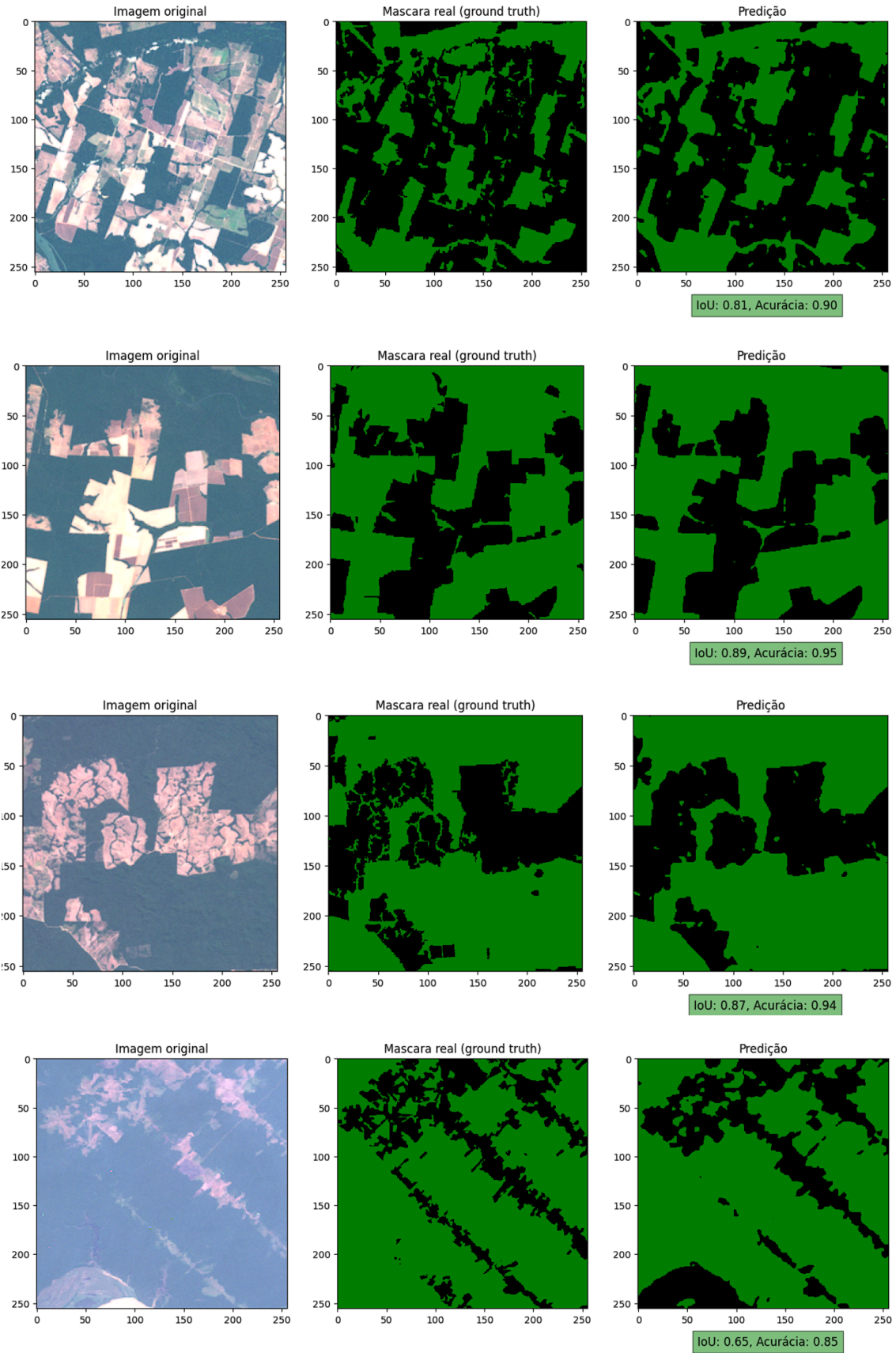


Figura 45: Segmentação Semântica. Amazonia 1. Fonte: (AUTOR, 2024).

Os resultados obtidos ressaltam a importância da resolução das imagens na precisão da segmentação. Através da Tabela 5, é possível visualizar o valor máximo de acurácia obtido em cinco trabalhos relacionados que utilizam redes neurais convolucionais, com uma média de 0.9472. O Sentinel-2, com sua alta resolução, foi capaz de fornecer detalhes essenciais para a identificação precisa de áreas menores de desmatamento, apresentando valores máximos de IoU e acurácia sendo 0.90748 e 0.97088 respectivamente, ficando acima da média entre as literaturas relacionadas. Por outro lado, o Amazonia 1, com sua resolução mais baixa, apresentou resultados inferiores à média, com valores máximos de IoU e acurácia de 0.81443 e 0.91177, respectivamente. Dessa forma, sua utilização se limita a aplicações mais abrangentes e menos precisas de monitoramento do desmatamento.

Projetos Relacionados	Satélite	Resolução	Acurácia
(AUTOR, 2024)	Sentinel-2	10m	0.97088
(AUTOR, 2024)	Amazonia 1	64m	0.91194
(MATOS, Carlos Henrique, 2024)	Sentinel-2	10m	0.9862
(ORTEGA ADARME, Mabel et al, 2020)	Landsat 8	30m	0.98
(Da Silva, Edson, 2023)	Sar Iceye	3m	0.96
(OLIVEIRA, Patrick Ágton, 2020)	Sentinel-2	10m	0.95
(SANTOS, Pedro, 2021)	RapidEye	30m	0.86

Tabela 5: Resultados de trabalhos relacionados (AUTOR, 2024).

5 CONCLUSÃO

O trabalho realizado teve como objetivo buscar uma nova metodologia capaz de detectar de forma automática áreas de desmatamento na Amazônia utilizando redes neurais artificiais convolucionais. O projeto visou alcançar a melhor eficácia na identificação de áreas desmatadas, utilizando uma tecnologia nacional acessível ao país. O foco foi maximizar o aproveitamento das imagens de satélite de baixa resolução disponíveis, e além disso, trazer uma comparação de resultados com um satélite de alta resolução.

Após a análise dos satélites Sentinel-2 e Amazonia 1, este trabalho expôs a metodologia atual de detecção de desmatamento implementada pelos programas Deter e Prodes, e introduziu uma nova abordagem utilizando inteligência artificial e a utilização de um satélite nacional para a detecção automática de áreas de desmatamento. Os modelos de segmentação semântica, empregando a arquitetura U-net, demonstraram-se eficazes na identificação de áreas desmatadas, representando uma ferramenta valiosa para reforçar as políticas de conservação ambiental.

O desafio principal enfrentado na realização do projeto foi a ausência de um dataset pronto contendo imagens e máscaras de segmentação para o satélite Amazonia 1. Para contornar essa limitação, foi necessário o desenvolvimento de um algoritmo capaz de extrair e converter as máscaras de desmatamento a partir do arquivo shapefile fornecido pelo INPE. Deste modo, a utilização de banco de dados e metodologia diferentes para obtenção das imagens entre os dois satélites pode ter sido um fator diferencial entre os resultados finais.

O intuito inicial do trabalho era tentar alcançar os resultados de segmentação semântica mais próximos possíveis entre os dois satélites Sentinel-2 e Amazonia 1. No entanto, nota-se que esse feito não foi possível devido às diferenças de resoluções espaciais. Esse desfecho já era esperado desde o início, porém ele destaca uma oportunidade para explorar métodos de aprimoramento de imagem que possam nivelar o campo de desempenho entre os dois satélites. Um exemplo promissor é o uso de Redes Adversárias Generativas (GANs) para super-resolução, que pode melhorar a qualidade das imagens de baixa resolução e, conseqüentemente, os resultados da segmentação.

Outra oportunidade futura visualizada através do trabalho realizado é a expansão do projeto para monitoração de outros biomas brasileiros, tais como o Cerrado, Caatinga, Pantanal, Mata Atlântica e o Pampa. Além disso, realizar a classificação do tipo de desmatamento ocorrido na região. Isso inclui classificar o tipo de desmatamento conforme metodologia PRODES e DETER, sendo eles corte seletivo, cicatriz de queimada, degradação,

mineração, etc. Tais abordagens seriam capazes de enriquecer o projeto e colaborar para para políticas de gestão e conservação mais efetivas.

REFERÊNCIAS

- Agência Espacial Europeia. “Sentinel-2 - Missions - Sentinel Online - Sentinel Online.” Sentinel Online, <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2>. Acesso em 9 de Abril de 2024.
- ANDRADE, Renan Bides de et al. Segmentação Semântica de Imagens Aplicada à Detecção de Desmatamento na Amazônia. 2020.
- BRAGAGNOLO, GRZYBOWSK, et al. “Amazon and Atlantic Forest image datasets for semantic segmentation.” Fevereiro 2021, <https://zenodo.org/records/4498086>.
- ARAGÓN, Luis Eduardo. A dimensão internacional da Amazônia: um aporte para sua interpretação/The international dimension of the Amazon: a contribution for its interpretation. Revista Nera, n. 42, p. 14-33, 2018.
- CAMPOS, Júlia Miranda Rossa. Avaliação de desempenho e de satisfação do usuário do assistente virtual Ifes. talk. 2022.
- CBS News. (2013, Fevereiro). NASA launches \$855 million Landsat mission. Retrieved from CBS News: <https://www.cbsnews.com/news/nasa-launches-855-million-landsat-mission/>
- CUNHA, Leonardo Cardoso da. Redes neurais convolucionais e segmentação de imagens: uma revisão bibliográfica. 2020.
- CUNHA, Leonardo Cardoso da. Redes neurais convolucionais e segmentação de imagens: uma revisão bibliográfica. 2020. 51 f. Monografia (Graduação em Engenharia de Controle e Automação) - Escola de Minas, Universidade Federal de Ouro Preto, Ouro Preto, 2020.
- DA COSTA PORTELLA, Diego Augusto Pereira et al. A importância da Amazônia na dinâmica climática do centro-sul brasileiro: Influência nas dinâmicas ambientais e socioeconômicas. Ensaios de Geografia, v. 9, n. 19, p. 66-86, 2022.
- DA SILVA, Edson Marques; SANTOS, Reginaldo. Uma arquitetura de rede neural convolucional para detecção de alvos de desmatamento na Amazônia através de imagens SAR ICEYE. In: Anais do XIV Congresso Brasileiro de Agroinformática. SBC, 2023. p. 230-237.
- DE ALMEIDA, Claudio Aparecido et al. METODOLOGIA UTILIZADA NOS SISTEMAS PRODES E DETER-2 a EDIÇÃO (ATUALIZADA). CEP, v. 12, p. 010, 2022.
- DE MEDEIROS, Talles Henrique et al. Computação Bioinspirada Aplicada à Robótica. 2005.
- DINIZ, Cesar Guerreiro et al. DETER-B: The new Amazon near real-time deforestation detection system. IEEE Journal of selected topics in applied earth observations and remote

sensing, v. 8, n. 7, p. 3619-3628, 2015.

DUMOULIN, Vincent; VISIN, Francesco. A guide to convolution arithmetic for deep learning. arXiv preprint arXiv:1603.07285, 2016.

Elian, Samir. “Meu teste deu positivo. E agora? Entendendo a sensibilidade e a especificidade dos testes diagnósticos - Meio de Cultura.” Blogs Unicamp, 16 June 2020, <https://www.blogs.unicamp.br/meiodecultura/2020/06/16/meu-teste-de-positivo-e-agora/>. Acesso 29 Abril 2024.

Empresas fornecedoras de imagens de satélite e a Capacidade de Dissuasão: uma alternativa. 2022.

FALCÃO, João Vitor Regis et al. Redes neurais deep learning com tensorflow. RE3C-Revista Eletrônica Científica de Ciência da Computação, v. 14, n. 1, 2019.

FERREIRA, Alessandro dos Santos. Redes neurais convolucionais profundas na detecção de plantas daninhas em lavouras de soja. 2017.

FLECK, Leandro et al. Redes neurais artificiais: Princípios básicos. Revista Eletrônica Científica Inovação e Tecnologia, v. 1, n. 13, p. 47-57, 2016.

Gao, Fengli, and Huicai Zhong. "Study on the Large Batch Size Training of Neural Networks Based on the Second Order Gradient. arXiv 2020." arXiv preprint arXiv:2012.08795.

GLOROT, Xavier; BENGIO, Yoshua. Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the thirteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, 2010. p. 249-256.

HE, Kaiming et al. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE international conference on computer vision. 2015. p. 1026-1034.

Grando, Nei. “A Essência do Aprendizado de Máquina.” Blog do Nei, 4 May 2022, <https://neigrando.com/2022/05/04/a-essencia-do-aprendizado-de-maquina/>. Accessed 25 April 2024.

IBGE. “Amazônia Legal 2022.” Amazônia Legal 2022, 2022, <https://www.ibge.gov.br/geociencias/cartas-e-mapas/mapas-regionais/15819-amazonia-legal.html?edicao=37404&t=o-que-e>. Acesso em 20 de Dezembro de 2023.

INPE. “Metodologia Utilizada Nos Sistemas Prodes e Deter - 2a Edição (Atualizada).” 2022, https://www.researchgate.net/publication/363153012_METODOLOGIA_UTILIZADA_NOS

_SISTEMAS_PRODES_E_DETER_-2_a_EDICAO_ATUALIZADA_INPE_Sao_Jose_dos_Campos_2022. Acessado 21 Abril 2024.

INPE. “PRODES — Monitoramento do Desmatamento da Floresta Amazônica Brasileira por Satélite.” INPE, <http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/prodes>. Acesso 29 Dezembro 2023.

IPCC, 2021: Sumário para Formuladores de Políticas. Em: Mudança do Clima 2021: A Base da Ciência Física. Contribuição do Grupo de Trabalho I ao Sexto Relatório de Avaliação do Painel Intergovernamental sobre Mudanças Climáticas [Masson-Delmotte, V., P. Zhai, A. Pirani, S. L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T. K. Maycock, T. Waterfield, O. Yelekçi, R. Yu and B. Zhou (eds.)]. Cambridge University Press. No Prelo.

JANIESCH, Christian; ZSCHECH, Patrick; HEINRICH, Kai. Machine learning and deep learning. *Electronic Markets*, v. 31, n. 3, p. 685-695, 2021.

JURASZEK, Guilherme Defreitas et al. Reconhecimento de produtos por imagem utilizando palavras visuais e redes neurais convolucionais. 2014.

KARRAS, Tero; LAINÉ, Samuli; AILA, Timo. A style-based generator architecture for generative adversarial networks. In: *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019. p. 4401-4410.

LEDIG, Christian et al. Photo-realistic single image super-resolution using a generative adversarial network. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017. p. 4681-4690.

MAGRINELLI, Marcelo; ANDRADE, Mateus Barros de. Empresas fornecedoras de imagens de satélite e a Capacidade de Dissuasão: uma alternativa. 2022.

MARIZ, Lucas de Freitas. *Redes geradoras adversárias em geração de imagens*. 2018.

MATOS, Carlos Henrique. Um estudo comparativo de desempenho de redes neurais da família Unet na tarefa de segmentação semântica de áreas de desmatamento na floresta amazônica. 2024.

MENESES, Paulo Roberto; ALMEIDA, T. de. *Introdução ao processamento de imagens de sensoriamento remoto*. Universidade de Brasília, Brasília, 2012.

MATOS, Carlos Henrique. Um estudo comparativo de desempenho de redes neurais da família Unet na tarefa de segmentação semântica de áreas de desmatamento na floresta amazônica. 2024.

Ministério da Ciência, Tecnologia e Inovação. “Imagens do satélite Amazonia-1 estão disponíveis para o público.” 03 11 2022, <https://www.gov.br/mcti/pt-br/acompanhe-o-mcti/noticias/2021/07/imagens-do-satelite-amazonia-1-estao-disponiveis-para-o-publico>. Accessed 09 04 2024.

Monteiro, Rosangela, et al. “INTELIGÊNCIA ARTIFICIAL, DEEP LEARNING, MACHINE LEARNING, REDES NEURAIS NA MEDICINA E BIOMARCADORES VOCAIS: CONCEITOS, ONDE ESTAMOS E PARA ONDE VAMOS.” 2022, https://socesp.org.br/revista/assets/upload/revista/17344834141648473754pdfpt01_revistasocesp_v32_01.pdf. Accessed 25 Abril 2024.

NEVES, Bárbara Coelho. Inteligência artificial e computação cognitiva em unidades de informação: conceitos e experiências. *Logeion: filosofia da informação*, v. 7, n. 1, p. 186-205, 2020.

Nobre AD. O Futuro Climático da Amazônia, Relatório de Avaliação Científica. Patrocinado por ARA, CCST-INPE, e INPA. São José dos Campos, Brasil, 42p, 2014.

OLIVEIRA, Patrick Ágton de. Deep learning na segmentação automática de imagens de satélite. 2020.

OLIVEIRA, Walysson Carlos dos Santos et al. Segmentação semântica de áreas de plantações agrícolas via U-Net em dois estágios. 2022.

ORTEGA ADARME, Mabel et al. Evaluation of deep learning techniques for deforestation detection in the Brazilian Amazon and cerrado biomes from remote sensing imagery. *Remote Sensing*, v. 12, n. 6, p. 910, 2020.

OSÓRIO, Fernando S.; BITTENCOURT, João Ricardo; OSÓRIO, Fernando Santos. Sistemas inteligentes baseados em redes neurais artificiais aplicados ao processamento de imagens. In: *I Workshop de inteligência artificial*. sn, 2000.

PACHECO, César Augusto Rodrigues. Deep learning conceitos e utilização nas diversas áreas do conhecimento. 2018.

PAIXÃO, Gabriela Miana de Mattos et al. Machine Learning na Medicina: Revisão e Aplicabilidade. *Arquivos brasileiros de cardiologia*, v. 118, p. 95-102, 2022.

PHIL, Kim. *Matlab deep learning with machine learning, neural networks and artificial intelligence*. Apress, New York, 2017.

Programa das Nações Unidas para o Meio Ambiente (PNUMA), et al. PERSPECTIVAS DO MEIO AMBIENTE NA AMAZÔNIA GEO AMAZÔNIA. 2008, https://antigo.mma.gov.br/estruturas/PZEE/_arquivos/geoamazonia_28.pdf.

Revista Pesquisa Fapesp. (2021). As Estrelas da Inovação. Um Satélite Brasileiro, pp. 80-81.
RIVERO, Sérgio et al. Pecuária e desmatamento: uma análise das principais causas diretas do desmatamento na Amazônia. *Nova economia*, v. 19, p. 41-66, 2009.

RODRIGUES, Diego Alves. Deep Learning e redes neurais convolucionais: reconhecimento automático de caracteres em placas de licenciamento automotivo. 2018.

RONNEBERGER, Olaf; FISCHER, Philipp; BROX, Thomas. U-net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*. Springer International Publishing, 2015. p. 234-241.

SANTOS, Pedro Thiago Cutrim dos. Aprimoramento da detecção de áreas de garimpo na Região dos Tapajós através de Redes adversárias de super-resolução. 2021.

SILVA, Thiago Araújo da. Uma rede UNet modificada para segmentação e identificação de lesões pulmonares em tomografias. 2021.

Silva Junior, C.H.L.; Pessoa, A.C.M.; Carvalho, N.S. et al. The Brazilian Amazon deforestation rate in 2020 is the greatest of the decade. *Nat Ecol Evol* 5, 144–145 (2021).
<https://doi.org/10.1038/s41559-020-01368-x>

SILVA, Wayrone Klaiton Luiz; GRANDE, Eliana Tiba Gomes; DE OLIVEIRA, Daniela Cabral. Estudo do satélite brasileiro Amazonia-1 E de sua trajetória: Mapeamento Sistemático e Análise Documental dos Artefatos Históricos–Oficiais. *Research, Society and Development*, v. 11, n. 2, p. e29011225894-e29011225894, 2022.

SOUZA, Victor et al. Análise comparativa de redes neurais convolucionais no reconhecimento de cenas. *Anais do Computer on the Beach*, v. 11, n. 1, p. 419-426, 2020.

TAYE, Mohammad Mustafa. Understanding of machine learning with deep learning: architectures, workflow, applications and future directions. *Computers*, v. 12, n. 5, p. 91, 2023.

Terrabrasilis – Plataforma de dados geográficos, <https://terrabrasilis.dpi.inpe.br/>. Acesso 29 Abril 2024.

“Thom Lane.” Transposed Convolutions explained with... MS Excel!, Medium, 2 Novembro 2018,
<https://medium.com/apache-mxnet/transposed-convolutions-explained-with-ms-excel-52d13030c7e8>. Accessed 09 Janeiro 2014.

VALDIONES, Ana Paula et al. Desmatamento Ilegal na Amazônia e no Matopiba: falta

transparência e acesso à informação. Culabá: Instituto Centro de Vida. <https://www.icv.org.br/publicacao/desmatamento-ilegal-na-amazonia-e-no-matopiba-falta-transparencia-e-acesso-a-informacao>, 2021.

Vilela, Paulo Augusto Arantes. "Detecção de Desmatamento no Estado de Mato Grosso do Sul utilizando Segmentação Semântica em imagens bi-temporais dos satélites Landsat 8 e Sentinel 2." (2024).

WANG, Xintao et al. Esrgan: Enhanced super-resolution generative adversarial networks. In: Proceedings of the European conference on computer vision (ECCV) workshops. 2018. p. 0-0.

ZHANG, Aston et al. Dive into deep learning. Cambridge University Press, 2023.

ZHOU, Zongwei et al. Unet++: A nested u-net architecture for medical image segmentation. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4. Springer International Publishing, 2018. p. 3-11.