

METACLASSIFICAÇÃO E REDE NEURAL CONVOLUCIONAL PARA SEGMENTAÇÃO DE DESMATAMENTO DA MATA ATLÂNTICA

Gabriel Domingos Adriano¹ Guilherme Alves Elias² Merisandra Côrtes de Mattos³

Resumo: O desmatamento é uma questão ambiental crítica que impacta diretamente a biodiversidade e os serviços ecossistêmicos, estando diretamente relacionado às mudanças climáticas e a eventos catastróficos cada vez mais frequentes na atualidade. No Brasil, a Mata Atlântica está entre os biomas mais devastados ao longo da história, restando menos de 20% de sua cobertura de vegetação original. Diante desse cenário, esforços têm sido empregados por órgãos governamentais e não governamentais com o intuito de conter o avanço do desmatamento. A identificação e o mapeamento das áreas desmatadas desempenham um papel fundamental, permitindo a alocação mais eficiente de recursos e o direcionamento de medidas preventivas. Nesse contexto, este estudo propõe o desenvolvimento de dois modelos, um utilizando o metaclassificador Random Forest e o outro a rede neural convolucional UNet, voltados para a segmentação de áreas desmatadas da Mata Atlântica no estado de Santa Catarina, visando identificar o modelo mais robusto para lidar com essa problemática. Para isso, algumas etapas foram seguidas como montagem da base de imagens capturadas pelo satélite Landsat-8 no período de 2020 a 2022, bem como os dados anuais de desmatamento da Global Forest Watch, implementação de técnicas de pré-processamento de dados, desenvolvimento dos modelos de aprendizado de máquina e aplicação das métricas de qualidade para avaliação dos resultados. O modelo de rede neural convolucional UNet gerou os melhores resultados, atingindo F1-Score de 82,14% na segmentação de desmatamento no ano de 2021, contra 49,99% do metaclassificador Random Forest.

¹Curso de Ciência da Computação, Grupo de Pesquisa em Inteligência Artificial Aplicada, Universidade do Extremo Sul Catarinense (Unesc), Criciúma - Santa Catarina - Brasil. dadriano.gabriel@unesc.net

²Coorientador, Curso de Ciências Biológicas, Universidade do Extremo Sul Catarinense (Unesc), Criciúma - Santa Catarina - Brasil. guilherme@unesc.net.

³Orientadora, Curso de Ciência da Computação, Grupo de Pesquisa em Inteligência Artificial Aplicada, Universidade do Extremo Sul Catarinense (Unesc), Criciúma - Santa Catarina - Brasil. mem@unesc.net.

Palavras-chave: Aprendizado de Máquina; Rede Neural Convolutacional; UNet; Metaclassificador; Random Forest; Desmatamento.

ABSTRACT: Deforestation is a critical environmental issue that directly impacts biodiversity and ecosystem services, and is directly related to climate change and increasingly frequent catastrophic events. In Brazil, the Atlantic Forest is among the most devastated biomes throughout history, with less than 20% of its original vegetation cover remaining. In light of this scenario, efforts have been made by governmental and non-governmental organizations to curb deforestation. The identification and mapping of deforested areas play a fundamental role, allowing for more efficient resource allocation and the direction of preventive measures. In this context, this study proposes the development of two models, one using the Random Forest metaclassifier and the other the UNet convolutional neural network, aimed at segmenting deforested areas of the Atlantic Forest in the state of Santa Catarina, aiming to identify the most robust model to handle this problem. For this, several steps were followed such as assembling the image base captured by the Landsat-8 satellite in the period from 2020 to 2022, as well as the annual deforestation data from Global Forest Watch, implementation of data preprocessing techniques, development of machine learning models and application of quality metrics for evaluation of results. The UNet convolutional neural network model generated the best results, achieving an F1-Score of 82.14% in deforestation segmentation in the year 2021, against 49.99% of the Random Forest metaclassifier.

Keywords: Machine Learning; Convolutional Neural Network; UNet; Meta-classification; Random Forest; Deforestation.

1 INTRODUÇÃO

Desmatamento é a ação de cortar ou derrubar árvores e florestas em grande área de cobertura, sem replantá-las, constituindo-se em um problema crônico que vem acontecendo ao redor do mundo e em todos os tipos de florestas existentes, mas principalmente em florestas tropicais (CFRN,). Dentre os motivos para a prática do desmatamento, um dos principais é a necessidade de madeira, que é utilizada por diferentes setores da sociedade, como por exemplo, na construção civil, na produção de papel e como fonte de energia. Também, a necessidade de terra para a agricultura e pecuária, especialmente para a criação de gado, têm ocasionado o desmatamento (Roser, 2022).

Historicamente, em 1500, com a colonização, a existência do Pau-Brasil deu início a exploração da Mata Atlântica que foi sendo derrubada para a construção de cidades e fazendas. Como consequência, foi o bioma que mais sofreu durante os ciclos econômicos brasileiros (UNEP, 2018).

A Mata Atlântica é um dos seis biomas brasileiros e se estende da região nordeste do país até a região sul, contribuindo diretamente no sustento de aproximadamente $\frac{3}{4}$ da população brasileira ao prover serviços ecossistêmicos como comida, regulação do ciclo da água e equilíbrio climático. Além disso, esse bioma é reconhecido pela Constituição Federal como Patrimônio Nacional e pela *United Nations Educational, Scientific and Cultural Organization* (UNESCO) como Reserva Mundial da Biosfera. Dada a sua importância, a Mata Atlântica é protegida no Brasil pela Lei da Mata Atlântica de 2006, que estabelece a proteção integral das florestas remanescentes e proíbe o corte de vegetação nativa (UNEP, 2018).

A Mata Atlântica em função do seu rico ecossistema, composto por espécies únicas de plantas e animais, está constantemente ameaçada (Hancock,). Segundo dados divulgados pela ONG SOSMA e INPE (2023), entre os anos de 2021 e 2022 a Mata Atlântica perdeu 20.075 hectares de vegetação, o que corresponde a mais de 20 mil campos de futebol. Desse total, 4.383 hectares de mata foram perdidos somente na região sul do país. Como consequência, foram lançados aproximadamente 9,6 milhões de toneladas de CO₂ equivalente na atmosfera.

Uma medida eficaz para a redução do desmatamento é identificar regiões com maior ameaça e direcionar recursos à essas áreas. Para isso, são necessárias informações sobre localização, gravidade e provável evolução da ameaça. Identificar essas regiões pode ser possível por meio do aprendizado de máquina. (Ball et al., 2022).

O Aprendizado de Máquina (AM) é um ramo da Inteligência Artificial (IA) que utiliza algoritmos para sintetizar as relações entre dados e informações. Devido a sua capacidade em detectar relacionamentos subjacentes em grandes quantidades de dados, esses algoritmos conseguem resolver problemas de análises de *big data*, reconhecimento de padrões e evoluções de informações (Awad; Khanna, 2015).

Dentre os algoritmos de aprendizado de máquina, tem-se o Random Forest (RF) que é um metaclassificador que pode ser usado para desenvolver modelos preditivos. RF é um conjunto de árvores de regressão e classificação, que utilizam divisões binárias em variáveis preditoras nos

dados de entrada para determinar previsões como saída. O benefício em utilizá-lo como modelo preditivo está na sua capacidade de lidar com conjuntos de dados com grandes quantidades de variáveis preditivas (Speiser et al., 2019).

Outra técnica utilizada em aprendizado de máquina, são as Redes Neurais Artificiais (RNA). As RNA se tornaram uma das ferramentas mais utilizadas do Aprendizado Profundo (*Deep Learning*), sendo empregada nos mais diversos campos como reconhecimento facial, veículos autônomos e na medicina. Elas têm a habilidade de extrair características dos dados de entrada de forma autônoma. A arquitetura de uma RNA é inspirada na percepção visual humana e possui neurônios, kernel e funções de ativação que simulam os sinais elétricos do cérebro que são disparados ao próximo neurônio quando atingem determinado limite (Li et al., 2022).

Diante do exposto, esta pesquisa compreendeu o desenvolvimento de dois modelos de aprendizado de máquina, a fim de segmentar o desmatamento da Mata Atlântica ocorrido entre um ano e outro no território de Santa Catarina, Brasil. O objetivo da pesquisa consiste em analisar os modelos gerados, por meio do algoritmo metaclassificador Random Forest e da rede neural convolucional UNet, referentes ao desmatamento da Mata Atlântica no estado de Santa Catarina. Para isso, métricas de qualidade para avaliação dos modelos de aprendizado de máquina gerados foram aplicadas.

2 TRABALHOS CORRELATOS

Sboui, Saidi e Lakti (2023) desenvolveram um modelo de Aprendizado de Máquina para identificar o desmatamento relacionado a extração do óleo de Palma na província de Aceh, Indonésia. Os autores utilizaram imagens dos satélites Sentinel e Landsat, capturadas entre os anos de 1988 e 2020. Para classificar se o desmatamento estava ou não relacionado a extração do óleo de Palma, os autores adotaram diferentes critérios, como por exemplo: distância da área desmatada em relação a área de plantação de Palma, distância de estradas, de moinhos, de água, entre outros. Como modelo, utilizaram o algoritmo Random Forest, devido a sua capacidade em lidar com dados em grande escala. Segundo os autores, os resultados foram satisfatórios, obtendo acurácia de 82%, F1-score de 69%, precisão de 72% e recall de 66%.

Alzu'bi e Alsmadi (2022) conduziram um trabalho na Jordânia baseado em segmentação semântica para monitoramento de desmatamento

em imagens de satélite. Para isso, utilizaram 4.576 imagens do satélite Landsat-8, capturadas entre 2010 e 2020. Na fase de pré-processamento, os autores reduziram o tamanho das imagens para 256 x 256 x 3, normalizaram os valores dos pixels e dividiram os conjuntos em treino, teste e validação, mantendo as proporções de 70%, 20% e 10%, respectivamente. Para esse estudo, os autores utilizaram três modelos de redes neurais: VGG-16, ResNet50 e Mobilenet-v2. Os três modelos pré-treinados atingiram acurácia semelhante. Mobilenet-v2, VGG-16 e ResNet50 atingiram, respectivamente, 94,8%, 94,7% e 94,3% de acurácia.

Ball et al. (2022) implementaram quatro modelos preditivos utilizando redes neurais convolucionais afim de identificar padrões espaço-temporal de desmatamento na Amazônia peruana. Como base de dados de desmatamento, utilizaram o dataset da Global Forest Change, que reúne dados anuais e globais sobre perdas de áreas florestais. As imagens utilizadas foram capturadas por satélite entre os anos de 2014 e 2020 e em quatro bandas espectrais. No pré-processamento, realizaram a normalização dos pixels e a divisão dos dados em 80% para treino e 20% para teste. O intuito era prever se, quando apresentado aos modelos, um pixel de floresta permaneceria como floresta no ano seguinte ou transicionaria para a classe não-floresta. Para isso, desenvolveram quatro modelos de redes neurais convolucionais: o primeiro utilizando duas dimensões de dados, o segundo três dimensões, o terceiro e o quarto utilizando hibridismo de rede neural recorrente por meio da célula LSTM. Como resultado, os terceiro e quarto modelos obtiveram as melhores acurácias, atingindo 94% e 93%, respectivamente.

Bem et al. (2020) conduziram esse trabalho a fim de testar e analisar algoritmos de aprendizado profundo na detecção de desmatamento na Amazônia brasileira. Para isso, utilizaram três regiões ao longo das rodovias Transamazônica e Cuiabá-Santarem. As imagens utilizadas foram capturadas pelo satélite Landsat-8 nos anos de 2017, 2018 e 2019, enquanto os dados de desmatamento foram retirados da plataforma PRODES, mantida pelo INPE. Os autores utilizaram três redes neurais: U-Net, SharpMask e ResUnet. Além desses modelos, também utilizaram os algoritmos Random Forest (RF) e Multilayer Perceptron (MLP). Segundo os autores, os modelos de aprendizado profundo tiveram vantagem em relação ao RF e MLP. O modelo ResUnet obteve as melhores métricas de avaliação, com exceção da precisão. O algoritmo Random Forest obteve os piores resultados.

3 MATERIAIS E MÉTODOS

Esta pesquisa visa a segmentação de áreas de desmatamento da Mata Atlântica ocorridas no período de um ano na região de Santa Catarina, Brasil, por meio de imagens de sensoriamento remoto geradas pelo satélite Landsat-8 e dados de desmatamento anuais retirados da base de dados da Global Forest Watch. Para isso, adotou-se uma abordagem de treinamento bi-temporal, na qual as imagens de dois anos consecutivos são empilhadas juntamente com os dados de desmatamento do ano alvo. Para isso, foram utilizadas imagens dos anos de 2020, 2021 e 2022, divididas em dois cenários de treino/teste. No primeiro cenário, tendo o ano de 2021 como alvo, utilizou-se as imagens de 2020 e 2021. No segundo cenário, 2022 foi considerado como o ano alvo e o treinamento realizado com os dados de 2021 e 2022.

Para isso, foram desenvolvidos dois modelos de aprendizado de máquina, o primeiro utilizando o algoritmo metaclassificador Random Forest e o segundo utilizando a Rede Neural Convolucional UNet, a fim de analisar os modelos obtidos por meio de métricas de qualidade empregadas em modelos classificatórios.

Em relação a natureza, trata-se de uma pesquisa aplicada e de base tecnológica que tem como objetivo a análise dos resultados dos dois modelos de aprendizado de máquina gerados quando aplicados a problemática do desmatamento. Em relação aos objetivos, é uma pesquisa exploratória, tendo em vista a necessidade da coleta de dados para o treinamento e teste dos modelos propostos. Quanto aos procedimentos, trata-se de uma pesquisa experimental, tendo em vista os experimentos realizados, a análise e a comparação dos resultados.

As etapas de desenvolvimento desta pesquisa compreenderam: seleção e montagem do conjunto de dados, composto por imagens de sensoriamento remoto e dados anuais de desmatamento; divisão e pré-processamento dos dados; desenvolvimento, treinamento e teste dos dois modelos; e aplicação das métricas de qualidade para análise dos modelos gerados.

3.1 BASE DE DADOS

O conjunto de imagens criado para esta pesquisa foi composto por 42 imagens capturadas pelo satélite Landsat-8 e 28 máscaras de desmatamento.

3.1.1 Imagens

O satélite Landsat-8 opera com onze bandas do espectro eletromagnético e captura as imagens em uma resolução espacial de 30m. No entanto, para este trabalho, foram utilizadas 7 bandas, correspondentes a aerosol costeiro, azul, verde, vermelho, infravermelho próximo, infravermelho de ondas curtas 1 e infravermelho de ondas curtas 2. A seleção dessas bandas específicas foi feita com base na relevância para a análise da vegetação e das características do solo nas regiões estudadas.

Além das sete bandas selecionadas, foi calculado o Índice de Vegetação por Diferença Normalizada (NDVI - Normalized Difference Vegetation Index) para cada imagem, que foi utilizado como a oitava banda. O cálculo é feito por meio da diferença das bandas do infravermelho próximo e do vermelho, dividido pela soma das bandas do infravermelho próximo e do vermelho. Os resultados variam de -1 a +1, onde valores mais próximos de +1 representam vegetação saudável.

As imagens foram obtidas por meio da *Application Programming Interface* (API) do Google Earth Engine (GEE) no formato GeoTIFF. Para isso, foi utilizada a linguagem de programação Python e o ambiente de programação do Google Colaboratory.

O GEE armazena as imagens do Landsat-8 em formato de número digital, que representa a intensidade da radiação capturada pelo sensor. Entretanto, esse formato não considera alguns fatores como distância sol-terra, atmosfera, entre outros. Portanto, foi necessário fazer a conversão desses números digitais para valores de reflectância de superfície, por meio do seguinte cálculo:

$$R = (ND \times 2,75 \times 10^{-5}) + (-0,2)$$

onde R é o valor de reflectância de superfície, ND é o valor bruto do pixel, $2,75 \times 10^{-5}$ é o multiplicador de reflectância e $-0,2$ é o termo de adição de reflectância. Esses valores foram fornecidos pelo GEE para a correção das bandas 1 a 7 do Landsat-8.

Para este trabalho, foram consideradas 14 regiões de interesse, conforme Figura 1 localizadas no estado de Santa Catarina, Brasil. Para cada região, foram criadas três coleções de imagens: a primeira contendo as imagens capturadas no período de 01/01/2020 a 31/12/2020, a segunda do período de 01/01/2021 a 31/12/2021 e a terceira abrangendo o período de 01/01/2022 a 31/12/2022. Além disso, cada imagem possui uma má-

cara de qualidade de pixel disponibilizada pelo GEE. Essa máscara identifica regiões da imagem com presença de ruídos como nuvens, sombras, neve, entre outros. Utilizando essas máscaras, foi possível mitigar presença de nuvens nas imagens. Ao todo, foram criadas 42 coleções. Com o objetivo de obter apenas uma imagem por ano e por região, calculou-se a mediana dos pixels das imagens de cada coleção, além de ser realizada a normalização dos valores para a escala [0, 1]. O resultado final foi um conjunto de 14 imagens por ano por região, sendo cada imagem composta por 7 bandas espectrais e mais uma banda extra referente ao NDVI, totalizando 8 bandas.

Figura 1 . Regiões de interesse utilizadas para treinamento e teste dos modelos.



Fonte: Do autor (2024).

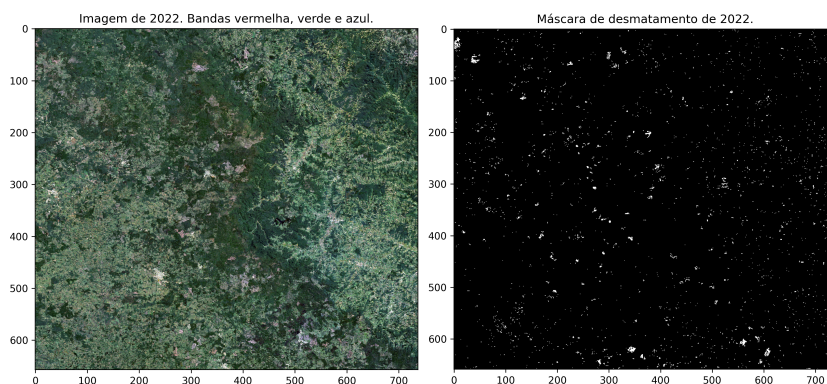
3.1.2 Máscaras de desmatamento

Para cada uma das 28 imagens geradas, referentes aos anos de 2021 e 2022, foi criada uma máscara binária correspondente. Tendo em vista que o ano de 2020 serviu apenas como base para o ano alvo de 2021, não foi necessário gerar uma máscara para esse período. Cada máscara destaca as áreas de desmatamento que ocorreram no ano específico representado pela imagem. Para realizar essa tarefa, foi utilizado o conjunto de dados "*Tree Cover Loss - Year of gross forest cover loss event (lossyear)*", disponível no repositório da Global Forest Watch, resultante de uma colaboração entre a Universidade de Maryland, Google, Serviço Geológico dos Estados Unidos e da *National Aeronautics and Space Administration* (NASA). Trata-se de uma base de dados global que registra a perda

florestal anual desde o ano de 2001 até 2022. Esses dados são disponibilizados em arquivos no formato GeoTIFF, sendo que o valor do pixel indica se houve ou não desmatamento e, em caso afirmativo, o ano em que o evento ocorreu.

Usando o ambiente Google Colaboratory e a linguagem Python, as máscaras originais de desmatamento foram processadas para se ajustarem às regiões de interesse. As máscaras foram redimensionadas para coincidir com a escala das imagens do Landsat-8 e filtradas de acordo com os anos de interesse, conforme exemplo na Figura 2, na qual têm-se uma imagem referente ao ano de 2022 de uma das regiões de interesse e sua respectiva máscara de desmatamento ocorrido no mesmo período. Esse processo resultou em 28 máscaras de desmatamento, cada uma correspondendo a uma das 14 regiões de interesse com os dados de desmatamento ocorridos no respectivo ano.

Figura 2. Imagem do Landsat-8 de uma região de Santa Catarina, seguida pela máscara de desmatamento de 2022 dessa mesma região.



Fonte: Do autor (2024).

3.2 MODELOS DE APRENDIZADO DE MÁQUINA GERADOS

Para essa pesquisa, foram desenvolvidos dois modelos de aprendizado de máquina: a rede neural convolucional UNet e o algoritmo metaclassificador Random Forest, por meio da linguagem de programação Python, do ambiente do Google Colaboratory Pro+⁴ e das bibliotecas Py-

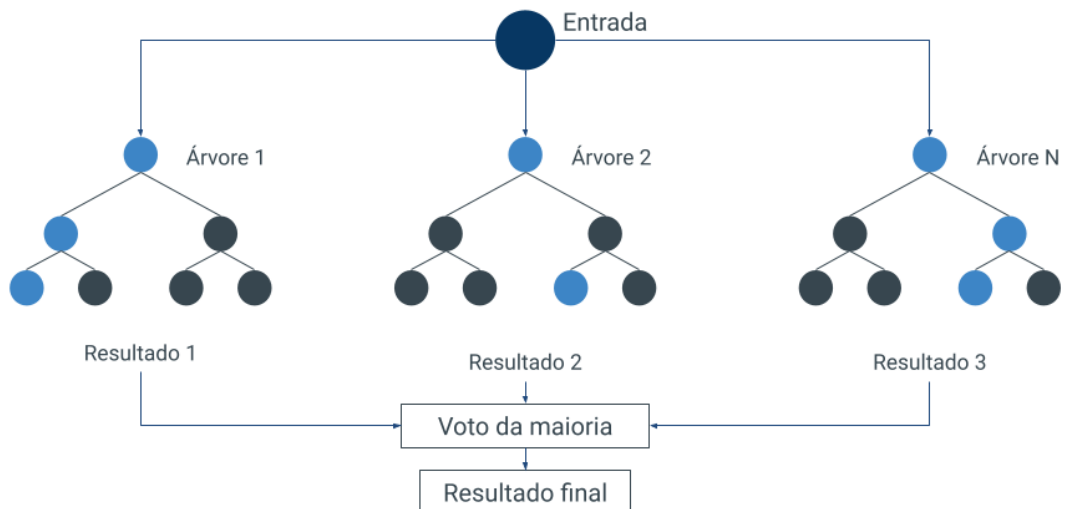
⁴Versão premium da plataforma Google Colaboratory.

Torch⁵ e Scikit-learn⁶.

3.2.1 Random Forest

De acordo com Breiman (2001), o algoritmo Random Forest (RF) consiste na combinação de múltiplas árvores de decisão, conforme Figura 3. Cada árvore é construída com base nos valores de um vetor aleatório, que é independente da amostragem e segue a mesma distribuição para todas as árvores da floresta. Seu poder reside no método de votação ou média, onde o resultado final é determinado pela decisão da maioria. Em problemas de classificação, a saída corresponde à classe mais votada. (Wang et al., 2018).

Figura 3. Algoritmo Random Forest.



Fonte: Do autor (2024).

Nesta pesquisa, foi utilizado o algoritmo Random Forest Classifier, disponibilizado pela biblioteca Scikit-learn. Esse algoritmo pode ser ajustado por diversos hiperparâmetros, como, por exemplo, o número de árvores de decisão e a profundidade máxima de cada árvore. Esses ajustes permitem otimizar o desempenho do modelo para diferentes tipos de dados e problemas, proporcionando maior flexibilidade e precisão nas previsões. Para cada cenário de teste, foram utilizados diferentes valores de hiperparâmetros, os quais serão detalhados na seção referente ao processo de

⁵PyTorch é uma biblioteca Python para aprendizado de máquina conhecida por sua flexibilidade na criação e treinamento de redes neurais artificiais.

⁶Scikit-learn é uma biblioteca de aprendizado de máquina em Python que oferece uma gama de algoritmos de aprendizado de máquina.

treinamento dos modelos.

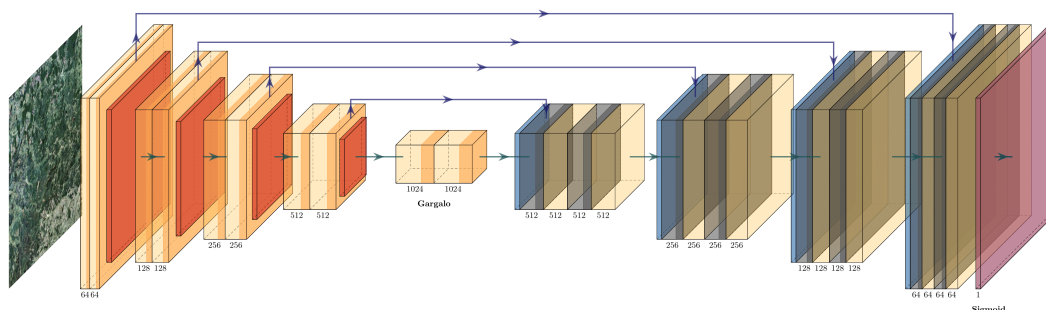
Considerando que o RF foi originalmente desenvolvido para lidar com dados tabulares, ele apresenta certas limitações em tarefas de segmentação de imagens. Uma dessas limitações é a incapacidade de levar em consideração a relação entre pixels adjacentes durante a análise. Para superar essa limitação, os dados de treinamento e teste foram redimensionados para o formato de vetor bidimensional, onde cada linha representa um pixel e cada coluna corresponde a uma característica (valor da banda de reflectância).

3.2.2 UNet

A UNet é uma rede neural convolucional profunda amplamente utilizada em problemas de segmentação de imagens. Originalmente desenvolvida para trabalhar com imagens biomédicas, nos últimos anos passou a ser utilizada em diferentes domínios de aplicação, incluindo segmentação em imagens de sensoriamento remoto (Wei et al., 2019).

Sua arquitetura é baseada em blocos de codificação e decodificação. O bloco de codificação é responsável por diminuir a dimensão espacial das imagens enquanto aumentam a profundidade das características, ao passo que o bloco de decodificação reconstrói a imagem para seu tamanho original (Wei et al., 2019).

Figura 4. Arquitetura da UNet.



Fonte: Do autor (2024).

Nesta pesquisa, foi desenvolvida uma arquitetura composta por três partes principais, conforme Figura 4: um bloco de codificação, um gargalo e um bloco de decodificação.

O bloco de codificação é composto por quatro blocos de convolução, seguidos por quatro camadas de *max pooling* a fim de reduzir

progressivamente o campo espacial da imagem. Cada bloco de convolução possui duas camadas convolucionais para extração de características, duas funções de normalização em lote para acelerar o treinamento e evitar possível *overfitting*⁷ e duas funções de ativação ReLu⁸ para introduzir não linearidade.

O bloco de gargalo é composto por uma camada de convolução adicional para capturar representações mais compactas da imagem de entrada.

O bloco de decodificação tem como objetivo reconstruir a imagem para sua dimensão espacial original. Ele consiste em quatro blocos de deconvolução, os quais aumentam progressivamente a resolução espacial da imagem. Na última etapa do bloco de decodificação, uma camada convolucional com um único canal de saída é aplicada para produzir a máscara de segmentação. Posteriormente, essa saída passa por uma função sigmoid para normalizar as probabilidades para um intervalo entre 0 e 1.

Além disso, foi utilizada a Focal Loss como função de perda devido ao desbalanceamento de classes. Para o ajuste dos pesos durante o treinamento, foi utilizado o otimizador Adam, tendo em vista sua eficiência computacional e taxa de aprendizado adaptativa. Para a taxa de aprendizado, foi utilizado um valor inicial de 0,001 por ser grande o suficiente para fazer progresso significativo e pequena o suficiente para não ignorar fatores importantes.

3.3 PROCESSO DE TREINAMENTO DOS MODELOS

Para o treinamento dos modelos, foi utilizada uma abordagem bi-temporal. Nesse método, para um ano alvo específico, o treinamento foi realizado empilhando as imagens de sensoriamento remoto do ano anterior e do ano alvo, resultando em uma única imagem composta por 16 bandas, considerando que cada imagem original possuía 8 bandas. Além disso, foram utilizadas as máscaras de desmatamento do ano alvo como rótulos. O objetivo dessa abordagem era permitir que os modelos captassem as modificações na vegetação ocorridas entre um ano e outro. Para este estudo, foram escolhidas duas etapas temporais: uma com o ano de 2021 como alvo, utilizando as imagens de 2020 e 2021, e outra com o ano de 2022 como alvo, utilizando as imagens de 2021 e 2022.

⁷Overfitting é quando um modelo de aprendizado de máquina se ajusta tão bem aos dados de treinamento que perde a capacidade de generalizar para novos dados.

⁸Função de ativação que transforma valores negativos em zero e mantém valores positivos inalterados.

3.3.1 Pré-processamento dos dados

Em tarefas de segmentação, cada pixel da imagem possui um rótulo que o identifica como pertencente a uma classe específica. Neste estudo, tem-se um problema de segmentação binária, pois os pixels são classificados em duas classes: não-desmatado e desmatado. O balanceamento das classes é crucial para se obter bons resultados. Sem esse balanceamento, o modelo tende a se sobreajustar à classe majoritária, resultando em um desempenho ruim na detecção da classe minoritária.

Ao se verificar a quantidade de amostras de pixel por classe, observou-se um expressivo desbalanceamento entre as classes, conforme demonstrado na Tabela 1.

Tabela 1. Dados das imagens e segmentos gerados, incluindo a distribuição de classes e o número total de pixels, antes e após o balanceamento.

Ano base/alvo	Nº imagens	Classe não-desmatado (px)	Classe desmatado (px)	Total (px)
Base original (32x32)				
2020/2021	139.593	141.748.507	1.194.725	142.943.232
2021/2022	139.593	141.768.588	1.174.644	142.943.232
Base original (64x64)				
2020/2021	34.383	139.656.578	1.176.190	140.832.768
2021/2022	34.383	139.680.764	1.152.004	140.832.768
Base balanceada (32x32)				
2020/2021	422	247.881	184.247	432.128
2021/2022	305	172.654	139.666	312.320
Base balanceada (64x64)				
2020/2021	188	593.870	176.178	770.048
2021/2022	150	472.787	141.613	432.128

Fonte: Do autor (2024).

Para resolver o desbalanceamento das classes, optou-se por uma estratégia de divisão das imagens originais em segmentos menores, e considerando somente os segmentos que contivessem, no mínimo, aproximadamente 30% da classe "desmatado". Essa porcentagem foi determinada empiricamente, levando em consideração a necessidade de uma base mais balanceada com a quantidade de amostras disponíveis. Embora uma porcentagem maior resultasse em uma base mais balanceada, ela também levaria a um número menor de amostras. Essa abordagem não apenas reduziu o desbalanceamento entre o número de amostras de cada classe, mas também diminuiu significativamente a quantidade total de imagens. Essa redução foi crucial para evitar sobrecarga computacional durante as fases de treinamento e teste dos modelos.

Para avaliar a abordagem mais eficaz na divisão dos segmen-

tos, foram empregadas duas estratégias distintas. Na primeira, as imagens foram divididas em segmentos de 32x32 pixels, com 10 pixels de sobreposição, considerando apenas segmentos contendo no mínimo 300 pixels da classe minoritária (classe "desmatado"). Na segunda estratégia, os segmentos foram criados com dimensões de 64x64 pixels e uma sobreposição de 20 pixels, sendo selecionados apenas os segmentos contendo no mínimo 600 pixels da classe "desmatado". Essa comparação permitiu identificar qual das duas abordagens oferece melhor desempenho na segmentação.

3.3.2 Cenários de Treino e de Teste

Com o propósito de avaliar os modelos desenvolvidos, foram elaborados seis cenários distintos (Tabela 2) de treinamento e de teste para cada algoritmo, variando os hiperparâmetros. Para cada cenário, o treinamento foi realizado duas vezes: uma utilizando imagens de tamanho 32x32 pixels e outra utilizando imagens de tamanho 64x64 pixels. Isso resultou em um total de 12 cenários distintos para cada algoritmo.

Tabela 2. Cenários de teste para os modelos desenvolvidos. Cada cenário foi executado duas vezes, alternando o tamanho dos segmentos das imagens.

Random Forest		
Cenário	Número de árvores	Profundidade máxima por árvore
1	500	None (máxima profundidade)
2	1000	None (máxima profundidade)
3	500	10
4	1000	10
5	500	30
6	1000	30

UNet		
Cenário	Tamanho do lote	Número de épocas
1	8	30
2	16	30
3	8	50
4	16	50
5	8	100
6	16	100

Fonte: Do autor (2024).

A primeira etapa foi a divisão da base de dados em subconjuntos. Para a rede neural convolucional, foram criados três subconjuntos denominados treino, validação e teste, com proporções respectivas de 70%, 10% e 20%. Para o Random Forest, que não requer uma base de validação separada, a divisão foi feita em dois subconjuntos denominados treino e teste, seguindo a proporção de 70% para treinamento, 20% para teste e

descartando os 10% restantes. Embora essas proporções não obedeçam a uma regra específica, são as mais utilizadas na literatura. Essa estratégia visa garantir o ajuste adequado dos modelos e evitar o *overfitting* durante o treinamento.

Todos os testes foram feitos utilizando a plataforma Google Colaboratory Pro+. Para o Random Forest, foi utilizado o ambiente de execução TPU V2⁹. Para a UNet, foi utilizado a A100 GPU⁹.

3.3.3 Métricas de Avaliação dos Modelos Gerados

Após a execução do treinamento e teste, foram aplicadas métricas de avaliação de modelos de aprendizado de máquina nos resultados das previsões, a fim de identificar o melhor modelo e o melhor cenário para a resolução do problema.

As métricas escolhidas para avaliar o modelo foram a acurácia, precisão, recall, F1-score e *Intersection over Union* (IoU). A acurácia indica a proporção total de previsões corretas feitas pelo modelo. A precisão revela a frequência com que, ao classificar uma amostra como positiva, a classificação está correta. O recall mostra a porcentagem de amostras verdadeiras que o modelo classificou corretamente como verdadeiras. O F1-score é a média harmônica entre a precisão e o recall, equilibrando as duas métricas para fornecer uma única medida. Por fim, a IoU mede a sobreposição entre duas regiões, sendo especialmente útil em tarefas de segmentação para avaliar a precisão das caixas delimitadoras e regiões segmentadas.

4 RESULTADOS E DISCUSSÕES

Após a aplicação das métricas avaliativas na etapa de teste dos modelos, os seguintes resultados foram gerados.

A Tabela 3 apresenta os resultados obtidos pelo algoritmo Random Forest quando treinado e testado utilizando imagens de tamanho 64x64. Ao analisar os resultados para ambos os períodos de treino e teste, e considerando que a classe *Desmatado* é mais importante para o problema, os cenários 3 e 4 mostraram os melhores resultados. Esses cenários demonstram valores mais altos de recall para a classe *Desmatado*, indicando uma maior capacidade de identificar corretamente as instâncias dessa classe. Além disso, o F1-Score é superior, refletindo um bom equilíbrio entre recall

⁹Ambiente de execução disponível no plano premium do Google Colaboratory. A plataforma pode ser acessada por meio do link: <https://colab.google/>

e precisão. Os valores de IoU também são mais elevados, demonstrando uma melhor sobreposição entre as máscaras reais e as geradas. Embora os outros cenários apresentem maior acurácia e melhores resultados para a classe *Não Desmatado*, os resultados para a classe *Desmatado* são inferiores. Considerando que o cenário 3 apresenta valores ligeiramente mais altos para o período 2021-2022 e tem um desempenho muito semelhante ao cenário 4 no período 2020-2021, o cenário 3 é a melhor escolha geral.

Tabela 3. Resultados dos cenários de treino e teste do algoritmo Random Forest, com imagens no tamanho 64x64 pixels.

Cenário	Classe	2020-2021					2021-2022				
		Precisão	Recall	F1	IoU	Acurácia	Precisão	Recall	F1	IoU	Acurácia
1	Não Desmatado	0,7789	0,9607	0,8603	0,0238	0,7570	0,7682	0,9718	0,8581	0,0143	0,7527
	Desmatado	0,2222	0,0395	0,0671			0,1888	0,0219	0,0392		
2	Não Desmatado	0,7789	0,9608	0,8603	0,0238	0,7570	0,7682	0,9719	0,8581	0,0143	0,7528
	Desmatado	0,2225	0,0395	0,0671			0,1889	0,0219	0,0392		
3	Não Desmatado	0,7851	0,4909	0,6041	0,1849	0,4989	0,7725	0,6032	0,6774	0,1697	0,5580
	Desmatado	0,2271	0,5268	0,3174			0,2352	0,4072	0,2982		
4	Não Desmatado	0,7853	0,4916	0,6047	0,1850	0,4993	0,7718	0,6029	0,6770	0,1689	0,5573
	Desmatado	0,2273	0,5266	0,3175			0,2343	0,4054	0,2970		
5	Não Desmatado	0,7787	0,9488	0,8554	0,0335	0,7501	0,7683	0,9684	0,8568	0,0177	0,7510
	Desmatado	0,2188	0,0505	0,0821			0,1949	0,0255	0,0451		
6	Não Desmatado	0,7787	0,9493	0,8556	0,0329	0,7504	0,7682	0,9686	0,8569	0,0174	0,7510
	Desmatado	0,2183	0,0499	0,0812			0,1938	0,0252	0,0445		

Fonte: Do autor (2024).

A Tabela 4 apresenta os resultados do modelo UNet para imagens 64x64. Considerando ambos os períodos, de treino e teste, e a maior importância da classe *Desmatado*, o Cenário 5 se destaca como a melhor escolha. Esse cenário mostrou altos valores de recall, precisão e F1-Score para a classe de interesse, especialmente no período 2021-2022, com um F1-Score de 0.7610 e recall de 0.7796. Embora outros cenários tenham apresentado boas métricas para a classe *Não Desmatado*, o desempenho superior do Cenário 5 na identificação correta da classe *Desmatado* o torna mais adequado para o problema.

Tabela 4. Resultados dos cenários de treino e teste do modelo UNet com imagens em 64x64 pixels.

Cenário	Classe	2020-2021					2021-2022				
		Precisão	Recall	F1	IoU	Acurácia	Precisão	Recall	F1	IoU	Acurácia
1	Não Desmatado	0,9025	0,9188	0,9106	0,5125	0,8600	0,9196	0,9245	0,9221	0,5888	0,8800
	Desmatado	0,6997	0,6561	0,6772			0,7457	0,7324	0,7390		
2	Não Desmatado	0,9249	0,8868	0,9055	0,5232	0,8563	0,9211	0,9278	0,9245	0,5971	0,8836
	Desmatado	0,6567	0,7505	0,7005			0,7554	0,7371	0,7461		
3	Não Desmatado	0,8998	0,9379	0,9185	0,5290	0,8708	0,9136	0,9200	0,9168	0,5600	0,8717
	Desmatado	0,7478	0,6381	0,6886			0,7290	0,7119	0,7203		
4	Não Desmatado	0,9072	0,9387	0,9227	0,5479	0,8779	0,9080	0,9254	0,9166	0,5472	0,8717
	Desmatado	0,7585	0,6671	0,7090			0,7363	0,6898	0,7123		
5	Não Desmatado	0,9034	0,9253	0,9142	0,5149	0,8652	0,9324	0,9187	0,9255	0,6015	0,8864
	Desmatado	0,7173	0,6570	0,6858			0,7433	0,7796	0,7610		
6	Não Desmatado	0,9017	0,9342	0,9177	0,5255	0,8699	0,9167	0,9208	0,9188	0,5688	0,8750
	Desmatado	0,7395	0,6471	0,6902			0,7340	0,7231	0,7286		

Fonte: Do autor (2024).

Além dos cenários testados com as imagens de tamanho 64x64, foram testados os mesmos cenários utilizando segmentos de 32x32. Segundo os resultados contidos na Tabela 5 para o modelo Random Forest, o cenário 4 se destacou como o mais eficaz para a classe *Desmatado* para ambos os períodos avaliados. Este cenário apresentou desempenho superior para as métricas precisão, recall, F1-Score e IoU, em relação aos outros cenários testados.

Tabela 5. Resultados dos cenários de treino e teste do modelo Random Forest com imagens em 32x32 pixels.

Cenário	Classe	2020-2021					2021-2022				
		Precisão	Recall	F1	IoU	Acurácia	Precisão	Recall	F1	IoU	Acurácia
1	Não Desmatado	0,5593	0,8278	0,6675	0,1336	0,5387	0,5665	0,6942	0,6239	0,2159	0,5302
	Desmatado	0,4397	0,1716	0,2469			0,4502	0,3204	0,3744		
2	Não Desmatado	0,5591	0,8327	0,6690	0,1297	0,5390	0,5663	0,6977	0,6252	0,2134	0,5304
	Desmatado	0,4389	0,1661	0,2410			0,4498	0,3163	0,3714		
3	Não Desmatado	0,5765	0,4829	0,5256	0,3245	0,5123	0,5792	0,4494	0,5061	0,3323	0,5077
	Desmatado	0,4556	0,5496	0,4982			0,4525	0,5822	0,5092		
4	Não Desmatado	0,5775	0,4813	0,5250	0,3260	0,5128	0,5795	0,4493	0,5062	0,3326	0,5079
	Desmatado	0,4563	0,5528	0,4999			0,4527	0,5828	0,5095		
5	Não Desmatado	0,5596	0,7660	0,6468	0,1743	0,5319	0,5657	0,6345	0,5981	0,2433	0,5214
	Desmatado	0,4412	0,2346	0,3063			0,4460	0,3766	0,4084		
6	Não Desmatado	0,5597	0,7721	0,6489	0,1706	0,5327	0,5656	0,6368	0,5991	0,2418	0,5215
	Desmatado	0,4413	0,2286	0,3012			0,4459	0,3740	0,4068		

Fonte: Do autor (2024).

Por fim, foi avaliado o modelo de rede neural convolucional UNet utilizando imagens de tamanho 32x32. Segundo os resultados descritos na Tabela 6, o cenário 5 se destaca como o mais eficaz, principalmente para a classe *Desmatado*. Este cenário apresentou maior F1-Score que os demais cenários, além de maior IoU para o período 2020-2021 e o segundo melhor para o período 2021-2022.

Tabela 6. Resultados dos cenários de treino e teste do modelo UNet com imagens em 32x32 pixels.

Cenário	Classe	2020-2021					2021-2022				
		Precisão	Recall	F1-Score	IoU	Acurácia	Precisão	Recall	F1	IoU	Acurácia
1	Não Desmatado	0.8705	0.8412	0.8556	0.6923	0.8393	0.8481	0.8510	0.8495	0.6544	0.8259
	Desmatado	0.8017	0.8368	0.8189			0.7954	0.7917	0.7936		
2	Não Desmatado	0.8549	0.8395	0.8471	0.6836	0.8285	0.8692	0.8288	0.8485	0.6713	0.8291
	Desmatado	0.7955	0.8143	0.8048			0.7800	0.8296	0.8040		
3	Não Desmatado	0.8590	0.8476	0.8532	0.6848	0.8350	0.8794	0.8099	0.8432	0.6753	0.8261
	Desmatado	0.8046	0.8185	0.8115			0.7655	0.8482	0.8047		
4	Não Desmatado	0.8711	0.8383	0.8544	0.6865	0.8383	0.8479	0.8353	0.8415	0.6475	0.8183
	Desmatado	0.7990	0.8383	0.8182			0.7794	0.7952	0.7872		
5	Não Desmatado	0.8768	0.8348	0.8553	0.6954	0.8401	0.8829	0.8060	0.8427	0.6730	0.8262
	Desmatado	0.7973	0.8471	0.8214			0.7631	0.8539	0.8059		
6	Não Desmatado	0.8564	0.8502	0.8533	0.6830	0.8345	0.8733	0.8113	0.8411	0.6664	0.8231
	Desmatado	0.8065	0.8141	0.8103			0.7649	0.8392	0.8003		

Fonte: Do autor (2024).

Conforme observado nos resultados apresentados, tem-se desempenho superior de ambos os modelos quando submetidos a imagens no tamanho 32x32, embora a UNet tenha tido um desempenho regular também nos cenários de 64x64, diferente do Random Forest. Isso se dá principalmente ao balanceamento das classes nesse cenário, conforme observado na Tabela 1, onde as classes apresentam um leve desbalanceamento nos segmentos de 64x64.

Além disso, quando comparados os desempenhos dos modelos Random Forest e UNet em seus melhores cenários, conforme Tabela 7, verifica-se um desempenho superior da rede neural convolucional. A arquitetura da UNet é projetada para analisar fatores espaciais da imagem, permitindo a captura de características contextuais e relacionamentos entre pixels. Em contraste, o Random Forest avalia pixels de maneira individual, sem considerar o contexto espacial ao redor de cada pixel, o que prejudica o modelo na tarefa de segmentação.

Tabela 7. Resultados dos modelos RF e UNet em seu melhor cenário.

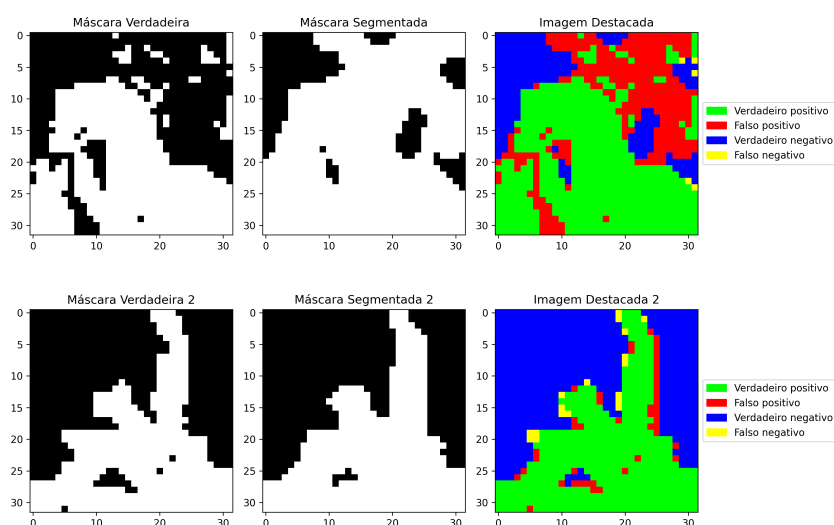
Modelo	Classe	2020-2021					2021-2022				
		Precisão	Recall	F1	IoU	Acurácia	Precisão	Recall	F1	IoU	Acurácia
RF	Não Desmatado	0,5775	0,4813	0,5250	0,3260	0,5128	0,5795	0,4493	0,5062	0,3326	0,5079
	Desmatado	0,4563	0,5528	0,4999			0,4527	0,5828	0,5095		
UNet	Não Desmatado	0,8768	0,8348	0,8553	0,6954	0,8401	0,8829	0,8060	0,8427	0,6730	0,8262
	Desmatado	0,7973	0,8471	0,8214			0,7631	0,8539	0,8059		

Fonte: Do autor (2024).

A fim de ilustrar os resultados da UNet em seu melhor cenário, a Figura 5 traz duas amostras aleatórias do conjunto de teste. A primeira imagem de cada amostra representa a máscara verdadeira de desmatamento ocorrido em algum local de Santa Catarina no ano de 2022. A segunda

imagem representa a máscara gerada pelo modelo UNet após analisar as imagens empilhadas de 2021 e 2022 dessa localidade. A terceira imagem representa a mesma máscara gerada pelo modelo, porém com identificações dos erros e acertos, sendo que pixel verde representa um verdadeiro positivo, pixel vermelho representa um falso positivo, pixel azul representa um verdadeiro negativo e pixel amarelo representa um falso negativo.

Figura 5. Exemplos de máscaras segmentadas pelo modelo UNet.



Fonte: Do autor (2024).

Considerando o melhor cenário de teste (Tabela 7), quando comparados os resultados obtidos pela rede neural convolucional UNet contra Sboui, Saidi e Lakti (2023), é possível constatar valores superiores nas métricas de precisão, recall e F1-Score. No entanto, quando comparados com Bem et al. (2020) e Ball et al. (2022), ambos realizados na Amazônia utilizando modelos de *deep learning*, os resultados foram inferiores. Essa diferença pode se dar devido a quantidade e qualidade dos dados utilizados. A Amazônia, sendo uma área de floresta densa, apresenta características que podem facilitar a detecção do desmatamento, resultando em dados mais definidos e modelos mais precisos.

5 CONCLUSÃO

Esta pesquisa teve como objetivo desenvolver dois modelos de aprendizado de máquina, um utilizando o metaclassificador Random Forest e o outro utilizando a rede neural convolucional UNet, a fim de segmentar

desmatamento da Mata Atlântica no estado de Santa Catarina, Brasil, ocorridos em 2021 e 2022, por meio de imagens de sensoriamento remoto, e comparar com métricas de qualidade os resultados gerados pelos modelos.

Após a seleção das imagens e máscaras de desmatamento, os modelos foram treinados e testados em 12 cenários diferentes, variando os hiperparâmetros e o tamanho das imagens. Métricas como F1-Score e IoU foram usadas para avaliar o desempenho dos modelos.

O modelo UNet obteve os melhores resultados em todas as métricas avaliadas e em todos os cenários testados, demonstrando vantagem do modelo de *deep learning* quando comparado com o metaclassificador Random Forest nesse domínio de aplicação.

O desbalanceamento significativo na base de dados original, bem como o número limitado de amostras após o balanceamento, podem ter impactado negativamente o desempenho dos modelos, especialmente o Random Forest.

Como possibilidades de trabalhos futuros sugere-se a comparação entre as redes neurais convolucionais UNet e Sharpmask na problemática de desmatamento, tendo em vista que ambas são focadas em resolver problemas de segmentação; ampliar o conjunto de dados para incluir outras regiões da Mata Atlântica em diferentes estados brasileiros, como Paraná e Rio Grande do Sul; explorar o desempenho dos modelos treinados com dados de outros biomas brasileiros, como a Amazônia, e testá-los na Mata Atlântica.

REFERÊNCIAS

ALZU'BI, A.; ALSMADI, L. Monitoring deforestation in Jordan using deep semantic segmentation with satellite imagery. **Ecological Informatics**, v. 70, p. 101745. ISSN 1574-9541. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1574954122001959>>.

AWAD, M.; KHANNA, R. **Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers**. 1st. ed. USA: Apress, 2015. ISBN 1430259892.

BALL, J. G. C. et al. Using deep convolutional neural networks to forecast spatial patterns of amazonian deforestation. **Methods in Ecology and Evolution**, v. 13, n. 11, p. 2622–2634. Disponível em: <<https://besjournals.onlinelibrary.wiley.com/doi/abs/10.1111/2041-210X.13953>>.

BEM, P. P. de et al. Change detection of deforestation in the Brazilian Amazon using Landsat data and convolutional neural networks.

Remote Sensing, v. 12, n. 6. ISSN 2072-4292. Disponível em: <https://www.mdpi.com/2072-4292/12/6/901>.

BREIMAN, L. Random forests. **Machine Learning**, Kluwer Academic Publishers, v. 45, n. 1, p. 5–32. ISSN 0885-6125. Disponível em: <http://dx.doi.org/10.1023/A%3A1010933404324>.

COALITION FOR RAINFOREST NATIONS. **Understanding Deforestation**. [S.l.]: <https://www.rainforestcoalition.org/understanding-deforestation/>. Acesso em: 10 jun. 2023.

HANCOCK, L. **What is the Atlantic Forest and why do we need to save it?** [S.l.]: <https://www.worldwildlife.org/stories/what-is-the-atlantic-forest-and-why-do-we-need-to-save-it>. Acesso em: 10 de junho de 2023.

LI, Z. et al. A survey of convolutional neural networks: Analysis, applications, and prospects. **IEEE Transactions on Neural Networks and Learning Systems**, v. 33, n. 12, p. 6999–7019.

ROSER, M. **What is a global forest transition and can it be achieved in our lifetime?** [S.l.]: <https://www.weforum.org/agenda/2022/05/humans-destroyed-forests-for-thousands-of-years-we-can-become-the-first-generation/>, 2022. Acesso em: 10 jun. 2023.

SBOUI, T.; SAIDI, S.; LAKTI, A. A machine-learning-based approach to predict deforestation related to oil palm: Conceptual framework and experimental evaluation. **Applied Sciences (Switzerland)**, MDPI, v. 13. ISSN 20763417.

SOS MATA ATLÂNTICA AND INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. **Atlas dos Remanescentes Florestais da Mata Atlântica: período 2021-2022**. 2023.

SPEISER, J. L. et al. A comparison of random forest variable selection methods for classification prediction modeling. **Expert Systems with Applications**, v. 134, p. 93–101. ISSN 0957-4174. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0957417419303574>.

UNITED NATIONS ENVIRONMENT PROGRAMME. **Municipalities in Brazil take the lead in conserving fragile Atlantic Forest biome**. 2018. Acesso em: 10 jun. 2023. Disponível em: <https://www.unep.org/news-and-stories/story/municipalities-brazil-take-lead-conserving-fragile-atlantic-forest-biome/>.

WANG, Y. et al. A novel consistent random forest framework: Bernoulli random forests. **IEEE Transactions on Neural Networks and Learning Systems**, v. 29, n. 8, p. 3510–3523.

WEI, S. et al. Multi-temporal sar data large-scale crop mapping based on u-net model. **Remote Sensing**, v. 11, n. 1. ISSN 2072-4292. Disponível em: <<https://www.mdpi.com/2072-4292/11/1/68>>.