

Aplicação e Otimização de Redes Neurais Convolucionais para Detecção de Desmatamento em Imagens de Satélite Sentinel-2

Leonardo de O. Tamasauskas¹, Jean A. C. Dias¹, Gabriel B. Costa³,
Marcos C. da R. Seruffo^{1,2}

¹ Laboratório de Pesquisa Operacional (LPO)
Universidade Federal do Pará (UFPA), Belém-PA

²Programa de Pós Graduação em Engenharia Elétrica
Universidade Federal do Pará (UFPA), Belém-PA

³Programa de Pós Graduação em Ciências Biológicas
Universidade Federal do Oeste do Pará (UFOPA), Santarém-PA

{jean.dias}@itec.ufpa.br,
{tamasauska,gabrielbritocosta}@gmail.com,
seruffo@ufpa.br

Abstract. *Monitoring deforestation in the Amazon is a complex challenge that requires the use of advanced technologies such as satellite image analysis and artificial intelligence (AI). Convolutional Neural Networks (CNNs) have been used to detect deforested areas in Sentinel-2 satellite images, with the U-Net architecture adjusted for forest area segmentation. To overcome the scarcity of labeled data, the Optuna algorithm was employed to optimize the model, improving its parameters and, consequently, enhancing the accuracy of the results.*

Resumo. *O monitoramento do desmatamento na Amazônia é um desafio complexo que exige o uso de tecnologias avançadas, como a análise de imagens de satélite e inteligência artificial (IA). Redes Neurais Convolucionais (CNNs) foram utilizadas para detectar áreas desmatadas em imagens do satélite Sentinel-2, com a arquitetura U-Net ajustada para segmentação de áreas florestais. Para superar a escassez de dados rotulados, o algoritmo Optuna foi empregado para otimizar o modelo, aprimorando seus parâmetros e, consequentemente, melhorando a acurácia dos resultados.*

1. Introdução

Monitorar o desmatamento na Amazônia é um desafio complexo que exige o uso de diversas ferramentas tecnológicas. Entre elas, a análise de imagens de satélite tem se destacado como uma abordagem eficaz para mapear as mudanças na cobertura florestal. No entanto, a tarefa de analisar cada imagem individualmente é extremamente demorada, especialmente considerando a vasta extensão da floresta amazônica. Para enfrentar essa limitação, a incorporação de inteligência artificial (IA) se tornou crucial, permitindo uma análise automatizada e eficiente de grandes volumes de dados [Pratiwi et al. 2021].

Neste contexto, a detecção de desmatamento baseada em redes neurais convolucionais (CNNs) se mostrou promissora devido à capacidade dessas redes em lidar com a análise de imagens de satélite. As CNNs são eficazes na extração de padrões

e características detalhadas que facilitam a identificação de mudanças no uso da terra, como o desmatamento. No entanto, a precisão desses modelos depende fortemente da quantidade de dados rotulados disponíveis para o treinamento, um desafio particularmente relevante em regiões como a Amazônia, onde a coleta e rotulação de dados são limitadas[Adarme et al. 2019].

Para enfrentar a limitação de dados rotulados, foi implementado o algoritmo de otimização de hiperparâmetros Optuna. Esse algoritmo é projetado para otimizar hiperparâmetros ao explorar de forma eficiente o espaço de possíveis configurações para um modelo. No contexto da detecção de desmatamento, Optuna permite encontrar as melhores combinações de hiperparâmetros, ajustando iterativamente o modelo para melhorar sua capacidade de identificar desmatamento com maior precisão. A aplicação do Optuna visa aprimorar os resultados do treinamento, possibilitando uma detecção mais precisa e robusta, mesmo com a quantidade limitada de dados rotulados disponíveis[Akiba et al. 2019].

2. Metodologia

A metodologia desta pesquisa foi organizada em três fases principais. Na primeira fase, foram coletadas e processadas imagens de satélite Sentinel-2 de três regiões no Estado de Mato Grosso, Brasil, resultando em 5.122 recortes de 256x256 pixels, com suas respectivas máscaras, que passaram por técnicas de pré-processamento para aumentar a variabilidade dos dados. Na segunda fase, um modelo de segmentação baseado na arquitetura U-Net foi desenvolvido por outro pesquisador, utilizando o MobileNetV2 como base de extração de características, para identificar áreas de desmatamento. Na fase final, foi utilizado o Optuna para otimizar os hiperparâmetros do modelo, ajustando fatores como o número de filtros e a taxa de dropout, com o objetivo de maximizar o coeficiente Dice e melhorar a precisão da segmentação. Esse processo garantiu uma abordagem eficiente para a detecção de desmatamento a partir de imagens de satélite.

2.1. Convolutional Neural Networks (CNNs) para Detecção de Desmatamento

Neste estudo, utilizamos Convolutional Neural Networks (CNNs), um tipo de arquitetura de deep learning amplamente utilizado para a análise de imagens, devido à sua capacidade de capturar padrões espaciais e hierárquicos em dados visuais. As CNNs são especialmente eficazes em tarefas de segmentação de imagens, onde o objetivo é identificar e delinear objetos dentro de uma imagem, tornando-as ideais para a detecção de desmatamento a partir de imagens de satélite.[Long et al. 2015]

2.2. Descrição do Modelo

Para a implementação da CNN, utilizamos o modelo U-Net, conhecido por sua eficácia em tarefas de segmentação semântica, particularmente em imagens médicas e sensoriamento remoto.[Carass et al. 2020] O U-Net foi adaptado neste trabalho para lidar com imagens de satélite de alta resolução, capturadas pelo Sentinel-2, um satélite da Agência Espacial Europeia (ESA) que oferece imagens multiespectrais com alta resolução espacial. Foram analisadas imagens de três diferentes regiões no Estado de Mato Grosso, Brasil, totalizando 5.122 recortes de imagens e máscaras de tamanho 256x256 pixels, derivadas de imagens sem nuvens.

A arquitetura do U-Net foi implementada utilizando o MobileNetV2 como a base de extração de características, aproveitando seus pesos pré-treinados. O MobileNetV2 é uma arquitetura leve e eficiente em termos de computação, projetada para dispositivos móveis, que preserva a precisão do modelo enquanto reduz o número de parâmetros e o tempo de inferência. Na nossa adaptação, utilizamos as saídas de camadas específicas do MobileNetV2 como entradas para o bloco de "up-sampling" do U-Net, que reconstrói a resolução original da imagem, enquanto mantém as características semânticas aprendidas pelas camadas de "down-sampling"[Dong et al. 2020].

2.2.1. Dados

O modelo foi treinado utilizando um banco de dados composto por imagens destinadas à segmentação semântica de áreas de floresta e não-floresta, empregando uma rede neural totalmente convolucional U-Net. O conjunto de treinamento incluiu 30 imagens no formato GeoTIFF, cada uma com resolução de 512x512 pixels, redimensionadas para 256x256 pixels para obedecer a entrada do modelo, acompanhadas por máscaras PNG correspondentes, onde as áreas de floresta são representadas em branco e as áreas de não-floresta em preto. Além disso, 15 imagens GeoTIFF e máscaras PNG associadas foram utilizadas para validar o desempenho do modelo U-Net durante o treinamento. Finalmente, o conjunto de teste, composto por 15 imagens GeoTIFF, foi empregado para avaliar a capacidade de generalização do modelo em novos dados.

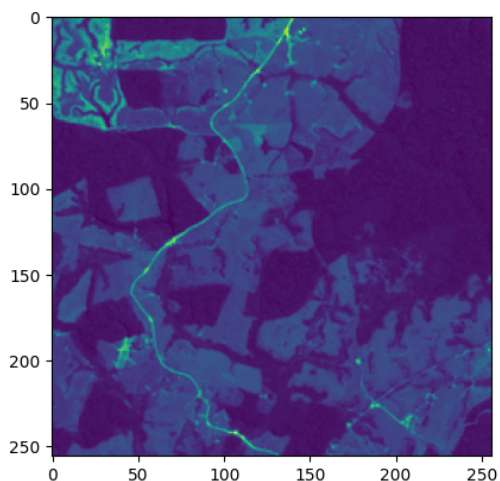


Figura 1. Imagem de Satélite

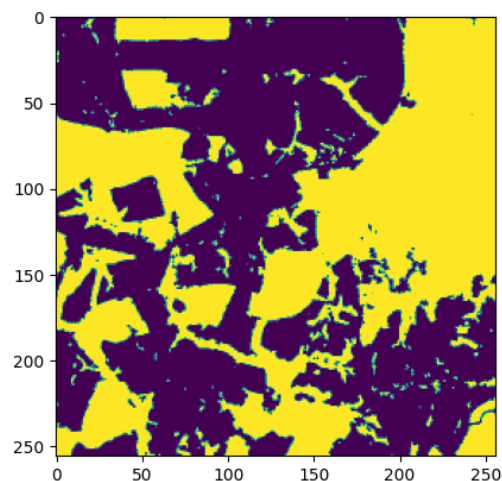


Figura 2. Imagem da Máscara

2.2.2. Otimização do Modelo com Optuna

Para maximizar a precisão do modelo, realizamos uma otimização de hiperparâmetros utilizando a biblioteca Optuna. Optuna é uma ferramenta que possui algoritmos que fazem otimização bayesiana, que permite a exploração automática e eficiente de espaços de hiperparâmetros, maximizando a performance do modelo[Akiba et al. 2019].

2.2.3. Coeficiente Dice

A função *dice_coef* calcula o Coeficiente Dice, que é uma métrica de avaliação frequentemente usada em tarefas de segmentação de imagens, como em modelos de redes neurais convolucionais (CNNs) para segmentação. A importância de otimizar o Coeficiente Dice está na sua capacidade de melhorar a qualidade das segmentações produzidas pelos modelos, o que tem um impacto direto em aplicações práticas onde a precisão é crítica [Shamir et al. 2019].

O Coeficiente Dice é dado por:

$$\text{Dice} = \frac{2 \times |A \cap B| + \text{Smooth}}{|A| + |B| + \text{Smooth}} \quad (1)$$

O parâmetro *smooth* é adicionado para evitar divisões por zero e melhorar a estabilidade numérica. O valor resultante é a média do Coeficiente Dice ao longo do eixo de lotes (batch).

2.2.4. Otimização

No então estudo, a função objetivo do Optuna foi definida para otimizar o Coeficiente Dice, por ser uma métrica amplamente reconhecida e eficaz para avaliar a qualidade da segmentação de imagens. A função objetivo ajustou dois hiperparâmetros principais: o número de filtros nas camadas convolucionais do U-Net e a taxa de dropout, que ajuda a prevenir o overfitting.

A rede foi treinada por 40 épocas, com o uso de callbacks para monitorar a performance durante o treinamento e interromper precocemente o processo caso o modelo não melhorasse após duas épocas consecutivas, garantindo assim a retenção dos melhores pesos do modelo. Após a conclusão dos 5 trials definidos para o estudo, o Optuna identificou os melhores parâmetros que maximizaram o Coeficiente Dice, contribuindo para a melhora nos resultados de treinamento do modelo.

3. Resultados

Durante o treinamento do modelo, o Coeficiente Dice foi utilizado como métrica principal para avaliar a sobreposição entre as previsões do modelo e os dados reais. Nas primeiras épocas, o modelo apresentou uma evolução significativa, com o *dice_coef* aumentando de 0.4744 na primeira época para 0.7660 na quinta, demonstrando uma melhora contínua no desempenho. No entanto, a partir da décima primeira época, o modelo começou a apresentar sinais de overfitting, com o *dice_coef* no conjunto de validação caindo, enquanto o desempenho nos dados de treino continuava a aumentar. Esses resultados indicaram a necessidade de ajustes para melhorar a generalização, como técnicas de regularização e otimização de hiperparâmetros.

O modelo otimizado pelo Optuna apresenta valores significativamente maiores de *dice_coef* tanto no treinamento quanto na validação. O melhor valor de *dice_coef* (0.9453) no modelo otimizado é superior ao melhor valor de *dice_coef* no modelo sem otimização (0.9057). O *val_dice_coef* também melhorou com a otimização, de 0.6906 para 0.7103.

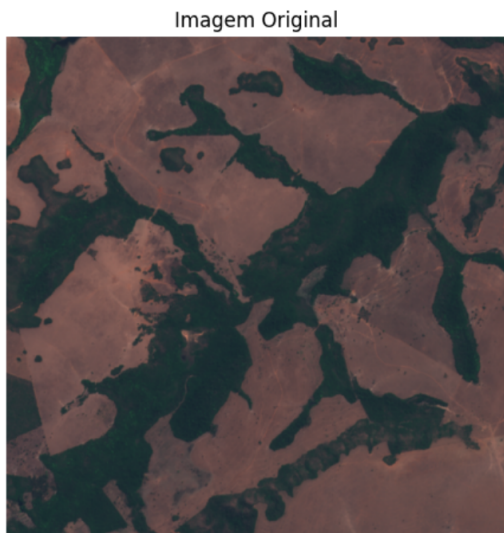


Figura 3. Imagem Original

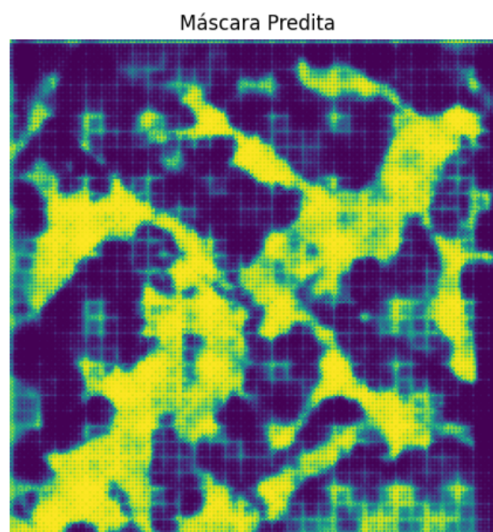


Figura 4. Máscara do Modelo Otimizado

A utilização do Optuna foi crucial para melhorar o desempenho do modelo. Optuna facilitou a otimização eficiente dos hiperparâmetros, resultando em um aumento significativo no Coeficiente Dice, de 0.9057 para 0.9453. Esse avanço ajudou a mitigar o overfitting e melhorou a capacidade do modelo de generalizar e detectar desmatamento com maior precisão. Assim, a otimização de hiperparâmetros revelou-se essencial para desenvolver um modelo mais eficaz.

Época	Coeficiente Dice (Antes da Otimização)	Coeficiente Dice (Após Otimização com Optuna)
1	0.4744	0.9384
6	0.7676	0.9453

Tabela 1. Comparação dos Valores de Coeficiente Dice Antes e Depois da Otimização com Optuna

4. Conclusão

O estudo demonstrou que a integração de inteligência artificial e técnicas de otimização de hiperparâmetros, como o Optuna, foi fundamental para melhorar a precisão e eficiência na detecção de desmatamento a partir de imagens de satélite, especialmente em regiões como a Amazônia, onde a coleta de dados é limitada. O uso de CNNs com a arquitetura U-Net provou ser eficaz para segmentação semântica, e a otimização dos hiperparâmetros permitiu maximizar o desempenho do modelo, mesmo com um conjunto de dados reduzido.

No entanto, a principal limitação do estudo foi justamente a falta de um grande número de dados, o que restringiu o potencial de generalização do modelo. Para mitigar esse desafio, estamos desenvolvendo alternativas para criar nossa própria base de dados, o que permitirá ampliar o escopo e a robustez das análises futuras. Ressaltamos que os resultados obtidos até o momento são de extrema importância, pois demonstram a viabi-

lidade e a eficácia das abordagens utilizadas, abrindo caminho para avanços significativos na detecção de desmatamento em regiões críticas como a Amazônia.

Referências

- [Adarme et al. 2019] Adarme, M. O., Happ, P. N., and Feitosa, R. Q. (2019). Assessment of an early fusion cnn approach applied to the deforestation detection in the brazilian amazon. In *XIX Simposio brasileiro de sensoriamento remoto, Sao Paulo, Brasil*, pages 1217–1220.
- [Akiba et al. 2019] Akiba, T., Sano, S., Yanase, T., Ohta, T., and Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.
- [Carass et al. 2020] Carass, A., Roy, S., Gherman, A., Reinhold, J. C., Jesson, A., Arbel, T., Maier, O., Handels, H., Ghafoorian, M., Platel, B., et al. (2020). Evaluating white matter lesion segmentations with refined sørensen-dice analysis. *Scientific reports*, 10(1):8242.
- [Dong et al. 2020] Dong, K., Zhou, C., Ruan, Y., and Li, Y. (2020). Mobilenetv2 model for image classification. In *2020 2nd International Conference on Information Technology and Computer Application (ITCA)*, pages 476–480. IEEE.
- [Long et al. 2015] Long, J., Shelhamer, E., and Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440.
- [Pratiwi et al. 2021] Pratiwi, N. K. C., Fu’adah, Y. N., and Edwar, E. (2021). Early detection of deforestation through satellite land geospatial images based on cnn architecture. *Jurnal Infotel*, 13(2):54–62.
- [Shamir et al. 2019] Shamir, R. R., Duchin, Y., Kim, J., Sapiro, G., and Harel, N. (2019). Continuous dice coefficient: a method for evaluating probabilistic segmentations. *arXiv preprint arXiv:1906.11031*.