

Semi-supervisionado para o desmatamento da Amazônia

Vitor Reis
UNIFESP (Universidade
Federal de São Paulo)
São José dos Campos, Brasil
vitor.reis@unifesp.br

Marcos Bauab
UNIFESP (Universidade
Federal de São Paulo)
São José dos Campos, Brasil
marcos.bauab@unifesp.br

Abstract— Este trabalho visa treinar um modelo para classificar imagens de satélite da Amazônia em “desmatado” e “floresta” através do algoritmo de aprendizagem semi-supervisionada NST (Noisy Student) e medir sua acurácia comparando com outros algoritmos.

O desmatamento na Amazônia é uma questão ambiental crítica, e há uma crescente necessidade de métodos precisos e eficientes para detectar e monitorar o desmatamento. Técnicas tradicionais de reconhecimento de imagem enfrentam dificuldades devido às características visuais complexas e diversas das imagens.

Os resultados mostram o potencial dessa abordagem em identificar atividades de desmatamento, auxiliando esforços de conservação e promovendo práticas sustentáveis na região. Pesquisas futuras podem se concentrar na otimização do algoritmo proposto e explorar sua aplicação em outros contextos ambientais.

I. INTRODUÇÃO

A Floresta Amazônica é uma das maiores e mais importantes florestas tropicais do mundo. Com uma extensão de aproximadamente 6,7 milhões de km², abrangendo nove países, incluindo o Brasil, a Amazônia é considerada um dos principais patrimônios naturais da humanidade e desempenha um papel vital na regulação do clima global. No entanto, nas últimas décadas, a região tem enfrentado um desafio significativo: o desmatamento.

O desmatamento na Amazônia representa uma ameaça direta à sustentabilidade ambiental, social e econômica da região e do mundo. Diante dessa problemática complexa e atual, a aplicação de técnicas de aprendizado de máquina pode

fornecer abordagens inovadoras e eficazes para monitorar e combater o desmatamento.

O monitoramento contínuo do desmatamento na Amazônia é uma tarefa desafiadora, considerando a vasta extensão territorial da floresta e a falta de recursos suficientes para realizar essa tarefa baseada em campo. Além do mais, a obtenção de dados rotulados, ou seja, dados que indicam com precisão as áreas desmatadas, pode ser um processo caro, demorado e sujeito a limitações logísticas.

Tendo isso em vista, os modelos semi-supervisionados surgem como uma solução promissora ao combinar dados rotulados e não rotulados, esses modelos exploram o potencial dos dados disponíveis para fornecer insights mais precisos e abrangentes sobre o desmatamento da Amazônia. Espera-se que um modelo semi-supervisionado possa identificar padrões e características relevantes associadas ao desmatamento, mesmo quando os dados rotulados são limitados.

II. CONCEITOS FUNDAMENTAIS

Um modelo de aprendizado semi-supervisionado é um tipo de modelo de aprendizado de máquina que utiliza uma combinação de dados rotulados e não rotulados para realizar tarefas de aprendizado.

Dados rotulados são amostras de dados que possuem rótulos ou etiquetas atribuídas a elas. Esses rótulos indicam a classe, categoria ou valor esperado associado a cada amostra. Por exemplo, em um problema de classificação de imagens de gatos e cachorros, os dados rotulados seriam as imagens

acompanhadas das informações de que cada imagem representa um gato ou um cachorro.

Dados não rotulados são amostras de dados que não possuem rótulos associados. Essas amostras não têm informações explícitas sobre a classe ou categoria a que pertencem. Em um conjunto de dados não rotulados, as amostras podem ser apenas observações em si, como imagens, texto ou dados numéricos, sem nenhuma informação sobre a classe ou categoria esperada.

Redes neurais são modelos computacionais inspirados no funcionamento do cérebro humano. Elas são compostas por um conjunto de unidades de processamento interconectadas, conhecidas como neurônios artificiais ou nós, que colaboram para realizar tarefas de aprendizado e processamento de informações.

O NST (Noisy Student Training) é um método de treinamento utilizado em modelos de deep learning para melhor desempenho do modelo em tarefas computacionais, como o reconhecimento de imagens.

A ideia do NST é usar um ciclo iterativo de treinamento e rotulagem de dados para aumentar a quantidade de dados de treinamento disponíveis para o modelo, quando não se tem uma base suficientemente grande. O processo pode envolver algumas etapas:

- 1. Treinamento de um modelo “professor”: é utilizado dados rotulados para treinar esse modelo. Esse processo gera dados sintéticos adicionais que são combinados com os dados originais para criar um conjunto de dados ampliado.
- 2. Treinamento de um modelo “estudante”: é utilizado o conjunto de dados ampliado para o treinamento do estudante que contém os dados gerados pelo modelo professor e os dados originais de treinamento, permitindo ampliar a base de dados com mais exemplos.
- 3. Repetição dos passos 1 e 2: são repetidos os passos 1 e 2, utilizando o estudante como novo professor, até encontrar um resultado que satisfaça o objetivo estabelecido inicialmente.

Acurácia balanceada é uma medida de desempenho que avalia a precisão do modelo na classificação de dados, ou seja, é uma forma de avaliação comumente usada para avaliar o quão bem um modelo está realizando uma tarefa. A acurácia é calculada dividindo o número de previsões corretas pelo número total de exemplos no conjunto de dados, já a acurácia balanceada calcula a média das taxas de acerto para cada classe, proporcionando uma medida mais equilibrada do desempenho geral do modelo.

Loss, também conhecida como função de perda ou função de custo, é uma medida que quantifica a discrepância entre as previsões feitas pelo modelo e os valores reais dos dados de treinamento. Em outras palavras, o “loss” mede o quão “errado” o modelo está sendo em suas previsões.

Existem diversas funções de perda que podem ser usadas dependendo do tipo de problema de aprendizado. Neste trabalho utilizamos o Categorical Cross-Entropy Loss que é utilizado em problemas de classificação entre duas classes, especialmente em combinação com ativação softmax na camada de saída da rede neural. Ela é calculada pela medida da diferença entre a distribuição de probabilidade prevista pelo modelo e a distribuição de probabilidade real dos rótulos.

A matriz de confusão é uma ferramenta usada para avaliar o desempenho de um modelo de classificação. Ela fornece uma representação das previsões feitas pelo modelo aos rótulos reais dos dados. Ela possui a seguinte estrutura:

- Verdadeiro positivo (TP): O modelo previu corretamente que a amostra pertence à "Classe A".
- Falso negativo (FN): O modelo previu incorretamente que a amostra pertence à "Classe B", quando na verdade ela pertence à "Classe A".
- Falso positivo (FP): O modelo previu incorretamente que a amostra pertence à "Classe A", quando na verdade ela pertence à "Classe B".
- Verdadeiro negativo (TN): O modelo previu corretamente que a amostra pertence à "Classe B".

O ruído Speckle adiciona variações multiplicativas aos pixels da imagem, simulando o efeito granulado. Para simular imagens de baixa qualidade, uma possível situação real.

III. TRABALHOS RELACIONADOS

Ao longo desse trecho, iremos abordar alguns trabalhos relacionados que serviram como inspiração e embasamento para o nosso. Cada um desses estudos contribuiu de forma significativa para o entendimento atual do tema, estabelecendo fundamentos teóricos valiosos.

Como primeiro trabalho a citar, devemos salientar o artigo “Self-training with Noisy Student improves ImageNet classification [1]” realizado por Qizhe Xie, Minh-Thang Luong, Eduard Hovy, Quoc V. Le, em conjunto com Google Research, Brain Team e Carnegie Mellon University.

Este trabalho nos trouxe a inspiração de utilizarmos a abordagem de aprendizado semi-supervisionado do NST (Noisy Student) como principal caminho para chegarmos ao resultado. O artigo demonstra ter alcançado resultados melhores do que o modelo state-of-the-art do ImageNet usando esta abordagem.

O algoritmo do Noisy Student foi usado no trabalho treinando um modelo professor com imagens rotuladas do ImageNet, e então usando esse modelo para rotular uma grande quantidade de imagens não rotuladas. A partir disso, treina-se um modelo de tamanho igual ou maior como o estudante, com uma

mistura de imagens rotuladas pelo professor e imagens do ImageNet com seus rótulos verdadeiros. Por fim, refaz-se quantas vezes achar necessário este mesmo procedimento, usando o estudante da etapa anterior como o novo professor. Em todo este processo, são adicionados ruídos para melhor generalizar os dados, ruídos esses tanto nos dados quanto no próprio modelo.

Dentre os trabalhos relacionados, destacamos o artigo “Evaluation of Deep Learning Techniques for Deforestation Detection in the Brazilian Amazon and Cerrado Biomes From Remote Sensing Imagery [2]” desenvolvido por Mabel Ortega, Raul Queiroz, Patrick Nigri, Claudio Aparecido e Alessandra Rodrigues, através do departamento de Engenharia Elétrica da PUC, Rio de Janeiro, Brasil, e Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, São Paulo, Brasil.

Esse artigo avalia técnicas baseadas em Deep Learning aplicadas à detecção de desmatamento em duas regiões tropicais com diferentes padrões de desmatamento: os biomas da Amazônia e Cerrado. Para comparações é utilizado um “SVM classifier” como linha de base. O principal objetivo é reduzir o esforço humano envolvido em programas de monitoramento, reduzindo custos e tempo de monitoramento de vegetação de biomas.

As principais contribuições deste trabalho são uma avaliação e comparação de três técnicas de Deep Learning para detecção automática de desmatamento nos biomas Amazônia brasileira e Cerrado; ou seja, Early Fusion (EF), Siamese Network (SN) e Convolutional SVM (CSVM), uma avaliação da precisão desses métodos em amostras de treinamento escassas e uma estimativa para cada método da relação: área designada como desmatamento versus área de desmatamento verdadeiro.

Podemos citar também o trabalho “Automatic Deforestation Detection based on the Deep Learning in Ukraine [3]” desenvolvido por Leonid Shumilo, Mykola Lavreniuk, Nataliia Kussul, Bella Shevchuk, através dos departamentos: National Technical University of Ukraine “Igor Sikorsky Kyiv Polytechnic Institute”, Kyiv, Ukraine, Department of Geographical Sciences, University of Maryland, College Park, MD, USA, Space Research Institute NASU-SSAU e Kyiv Academic University.

Este trabalho apresenta abordagens de Sensoriamento Remoto de Aprendizagem Profunda combinando imagens de radar de abertura óptica e sintética de satélites, nas quais a detecção de objetos é realizada por um sistema neural baseado em U-net, uma rede treinada com o uso de aprendizado semi-supervisionado. São testadas sua eficiência na região de Kiev para possível implementação nas demais regiões.

Outro interessante trabalho é o “Detecting Deforestation from Satellite Imagery [4]” desenvolvido por Nick Condo e publicado na “Towards Data Science”.

O trabalho tem como objetivo rastrear as mudanças na floresta amazônica devido ao desmatamento usando dados de satélite. A tarefa é um problema de classificação de imagens

multi-rótulo, onde cada imagem terá um e potencialmente mais de um rótulo atmosférico e zero ou mais rótulos comuns e raros.

O autor utiliza a biblioteca “fastai”, que simplifica o treinamento de redes neurais usando melhores práticas modernas permitindo que a taxa de aprendizado mude ciclicamente entre valores limite razoáveis ao invés de diminuir monotonicamente a taxa de aprendizado.

Ele também utiliza um método de redimensionamento progressivo que é um método de treinamento em que você primeiro treina o modelo em um tamanho menor que o original e continua aumentando o tamanho da imagem duas ou três vezes até finalmente chegar ao tamanho original.

No entanto, o autor diz que tentou a utilização de um modelo semi-supervisionado usando os rótulos do melhor envio para treinar no conjunto de teste, mas o desempenho do seu modelo diminuiu ligeiramente.

Outro trabalho que está relacionado com a detecção de desmatamento na floresta Amazônica é o “Detecting deforestation in the Amazon rainforest using unsupervised K-means clustering on satellite imagery [5]” disponível no ArcGIS Developers.

A sessão encontrada faz comentários e fornece um exemplo de modelo não supervisionado, modelo que não usaremos, utilizando o método K-means clustering, ou seja, usaria características das imagens para agrupá-las em grupos de acordo com suas semelhanças, ao qual também não usaremos.

Um trabalho que não está diretamente relacionado com o nosso tema mas trata de reconhecimento e recuperação de imagens é o “Reconhecimento e Recuperação de Imagens Utilizando Redes Neurais Artificiais do Tipo MLP [6]” desenvolvido por Túlio Cléber Bender e Fernando Santos Osório pela Universidade do Vale do Rio dos Sinos (UNISINOS).

O trabalho aplica técnicas de Inteligência Artificial (IA) para executar o reconhecimento e recuperação de imagens bitmap a partir de características extraídas automaticamente das imagens. A aplicação faz uso de técnicas de Machine Learning da IA com aprendizado supervisionado, onde se optou pelo uso de Redes Neurais Artificiais (RNA) do tipo Multilayer Perceptron (MLP) com algoritmo de aprendizado construtivo e incremental do tipo Cascade-Correlation [FAHLMAN 91].

Nessa mesma linha podemos citar o “Remote sensing for mineral exploration [7]” desenvolvido por Floyd F. Sabins.

Esse relatório comenta sobre o sensoriamento remoto, ciência de adquirir, processar e interpretar imagens e dados relacionados, adquiridos de aeronaves e satélites que registram a interação entre matéria e energia eletromagnética. Ele descreve os sistemas de sensoriamento remoto empregados na exploração mineral e apresenta as técnicas de computador

usadas para processar dados digitais adquiridos pelos sistemas e como dados multiespectrais são processados digitalmente para reconhecer minerais.

Temos também o “Consistency Analysis of Remote Sensing Land Cover Products in the Tropical Rainforest Climate Region: A Case Study of Indonesia [8]” desenvolvido por Junmei Kang ,Zhihua Wang, Lichun Sui ,Xiaomei Yang e Yuanzheng Ma Jun Wang

O artigo comenta um pouco da importância da aquisição de informações de cobertura da terra com base em sensoriamento remoto devido à sua capacidade de determinar rapidamente a distribuição da cobertura da superfície da terra e informações de mudanças dinâmicas na região de interesse e que ao mesmo tempo, não se deve ignorar que diferentes produtos podem usar diferentes sistemas de classificação, métodos de classificação e imagens de satélite. Portanto, os próprios produtos de cobertura da terra resultantes, bem como a série de estudos de acompanhamento derivados desses produtos de cobertura da terra, irão variar.

Um artigo que nos chamou atenção foi “ForestEyes Project - Citizen Science and Machine Learning to detect deforested areas in tropical forests [9]” desenvolvido por Fernanda B. J. R. Dallaqua*, Fabio A. Faria, e Álvaro L. Fazenda , pela Universidade Federal de São Paulo (Unifesp).

Este artigo comenta sobre o Projeto “ForestEyes”, ao qual voluntários analisam e classificam imagens de sensoriamento remoto, e esses dados são usados para treinar técnicas de aprendizado de máquina que detectarão o desmatamento em novas imagens de sensoriamento remoto. Neste programa, especialistas interpretam imagens do satélite Landsat-8, delimitando os polígonos de desmatamento e calculando a taxa anual de desmatamento em áreas da Amazônia Legal.

Foram usados nove classificadores diferentes como k-Vizinhos mais próximos (kNN), Perceptron multicamada (MLP), Random Forest (RF) e Support Vector Machines (SVM). Os resultados foram comparados com a Amostragem Aleatória (RS).

Existem diversos outros trabalhos que estão ligeiramente relacionados com o nosso que valem a pena ser citados como “Reconhecimento de imagens Bidimensionais utilizando Redes Neurais Artificiais [10] ” desenvolvido por Guy Perelmutter, Enrique Vinicio Carrera E., Marley Vellasco e Marco Aurélio Pacheco ou o “ Deforestation Detection with weak Supervised Convolutional Neural Networks in Tropical Biomes [11]” desenvolvido por P. J. Soto, G. A. O. P. Costa , M. X. Ortega , J. D. Bermudez , R. Q. Feitosa.

Além destes, um trabalho que deve ser citado é “Deforestation processes in south-western Madagascar over the past 40 years: what can we learn from settlement characteristics? [15]” realizado por Katja Brinkmann *, Fanambinantsoa Noromiarilanto, Rakotomalala Yedidya Ratovonamana e Andreas Buerkert.

Este artigo aborda a perda de cobertura florestal e mudanças no uso da terra em Mahafaly, sudoeste de Madagascar, ao longo de 40 anos. Utilizando imagens de satélite do Landsat e dados socioeconômicos coletados por meio de entrevistas, as imagens foram classificadas por um método supervisionado para identificar tendências de desmatamento e mudanças na cobertura da terra. Os assentamentos estudados foram classificados com base em sua localização geográfica, atividades socioeconômicas e status econômico.

IV. OBJETIVO

Avaliar a eficácia de um modelo semi-supervisionado aplicado ao desmatamento da Amazônia. Buscamos investigar como a combinação de dados rotulados e não rotulados pode melhorar a precisão e a capacidade de generalização do modelo, contribuindo para um monitoramento mais eficiente e preciso do desmatamento.

V. METODOLOGIA EXPERIMENTAL

A. BASE DE DADOS

A base de dados inicialmente era constituída de mais de 40 mil imagens feitas por satélite da floresta Amazônica, cada uma delas de 224x224 pixels. Todas elas possuíam classificações previamente estabelecidas, com mais de 449 tipos de classificações diferentes, sendo que cada imagem tinha uma quantidade variada de classes. Como exemplo segue uma imagem representativa:



Imagem de satélite com estradas e florestas

Para utilizarmos dessa base de dados em nosso trabalho, escolhemos utilizar uma quantidade menor de imagens, tomando 20 mil imagens e escolhendo 10 mil de forma

aleatória, dentre as imagens tomadas reduzimos o tamanho delas para a metade, facilitando assim o tratamento. Além disso, retiramos imagens que não seriam de grande valor para treinar o modelo escolhido, como imagens com muitas nuvens ou com difícil visualização e resumimos a quantidade de classes em apenas duas, “floresta” e “desmatado”. A retirada das imagens foi feita com base na classificação prévia delas. Como exemplo, temos a seguinte imagem, que demonstra a classe “cloudy” da base de dados:



Imagem com muitas nuvens, impedindo a visualização

B. PROTOCOLO DE VALIDAÇÃO

Como protocolo de validação utilizamos o K-Fold. O protocolo K-fold divide o conjunto de dados em K partições ou dobras (folds) de tamanhos aproximadamente iguais. O valor de K é escolhido antecipadamente e representa o número de vezes que o processo de treinamento e validação será repetido. O procedimento então é separado em:

1. Dividir os dados em K partições em que uma das partições é usada como validação em uma das iterações. Ou seja, se temos 5 folds, o primeiro pode ser usado como validação na primeira iteração e nas outras iterações será utilizado apenas como treinamento.
2. Repetir o treinamento e a validação K vezes e obter métricas de desempenho em cada iteração, como acurácia e a “loss function”.

Este protocolo favorece a utilização mais eficiente dos dados, especialmente quando a quantidade de dados é limitada, pois ajuda a reduzir a variância nos resultados de validação, dado que cada fold é usado em algum momento tanto para validação quanto para treinamento.

C. MÉTODOS

Utilizaremos a linguagem de programação Python, em conjunto com o Jupyter Notebook para executar os trechos de

código. Seguiremos a abordagem de aprendizado do Noisy Student, utilizando imagens do banco de imagens [14], filtraremos os rótulos que fazem sentido ao escopo do nosso trabalho, treinaremos o primeiro modelo, uma Rede Neural, professor com essas imagens rotuladas baseado no algoritmo do NST, a partir da rotulação do modelo professor treinaremos um modelo estudante, também uma Rede Neural, com adição de ruído nos dados e modelo, como *data augmentation* e *dropout*, e por fim realizaremos experimentos de resultados para chegar a um número de repetições, de estudante se tornando professor, satisfatório.

A partir dos resultados obtidos, faremos comparações com outros modelos da literatura ou com outras arquiteturas de Redes Neurais.

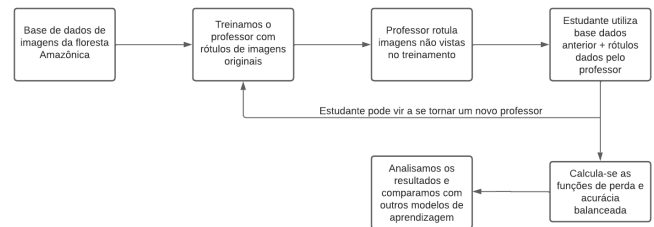


Diagrama de blocos do pipeline experimental

VI. O QUE SERÁ ENTREGUE NO FINAL?

Entregaremos um modelo treinado a partir de aprendizagem semi-supervisionada que seja capaz de inferir imagens de áreas desmatadas e áreas preservadas e separá-las.

REFERENCES

- [1] Self-training with Noisy Student improves ImageNet classification. Disponível em: <<https://paperswithcode.com/paper/self-training-with-noisy-student-improves>>. Acesso em: 6 jun. 2023.
- [2] ORTEGA ADARME, M. 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, 12 mar. 2020.
- [3] SHUMILO, L. et al. Automatic Deforestation Detection based on the Deep Learning in Ukraine. [s.l.: s.n.]. Disponível em: <<https://ela.kpi.ua/bitstream/123456789/46162/1/Automatic%20Deforestation%20Detection%20based.pdf>>. Acesso em: 4 jun. 2023.
- [4] CONDO, N. Detecting Deforestation from Satellite Imagery. Disponível em: <<https://towardsdatascience.com/understanding-the-amazon-rainforest-with-deep-learning-732bfb2eca6e>>.
- [5] Detecting deforestation in the Amazon rainforest using unsupervised K-means clustering on satellite imagery. Disponível em: <<https://developers.arcgis.com/python/samples/detecting-deforestation-using-kmeans-clustering-on-sentinel-imagery/>>. Acesso em: 4 jun. 2023.
- [6] BENDER, T.C.; OSÓRIO, F.S. Reconhecimento e Recuperação de Imagens Utilizando Redes Neurais Artificiais do Tipo MLP.
- [7] SABINS, F. F. Remote sensing for mineral exploration. Ore Geology Reviews, v. 14, n. 3-4, p. 157–183, set. 1999.

- [8] KANG, J. et al. Consistency Analysis of Remote Sensing Land Cover Products in the Tropical Rainforest Climate Region: A Case Study of Indonesia. *Remote Sensing*, v. 12, n. 9, p. 1410, 30 abr. 2020.
- [9] DALLAQUA, F.; FARIA, F.; FAZENDA, L. ForestEyes Project -Citizen Science and Machine Learning to detect deforested areas in tropical forests. [s.l.: s.n.]. Disponível em: <https://sol.sbc.org.br/index.php/sibgrapi_estendido/article/download/20008/19836/>. Acesso em: 4 jun. 2023.
- [10] PERELMUTER.;CARRERA.; MARLEY.; AURÉLIO P. Reconhecimento de Imagens Bidimensionais Utilizando Redes NeuraisArtificiais.Disponívelem:<http://sibgrapi.sid.inpe.br/col/sid.inpe.br/sibgrapi/2013/02.18.15.52/doc/25%20Reconhecimento%20de%20imagens%20bidimensionais.pdf>. Acesso em: 4 jun. 2023.
- [11] SOTO, P. J. et al. DEFORESTATION DETECTION WITH WEAK SUPERVISED CONVOLUTIONAL NEURAL NETWORKS IN TROPICAL BIOMES. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, v. XLIII-B3-2022, p. 713–719, 30 maio 2022.
- [12] CLEMENT, C. R.; HIGUCHI, N. A floresta amazônica e o futuro do Brasil. *Ciência e Cultura*, v. 58, n. 3, p. 44–49, 1 set. 2006.
- [13] RUSSELL, S. J. et al. *Inteligencia artificial : un enfoque moderno*. Madrid: Pearson Educación, D.L, 2011.
- [14] Banco de imagens disponível em: <https://www.kaggle.com/datasets/prosperchuks/amazonsatelliteimages>
- [15] Deforestation processes in south-western Madagascar over the past 40 years: what can we learn from settlement characteristics? Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0167880914003284?fr=RR-2&ref=pdf_download&rr=7db86ab22e876b18>. Acesso em 22 de junho de 2023.