APLICAÇÃO DE SELF-SAMPLING PARA SEGMENTAÇÃO DE ÁRVORES EM CENÁRIO DE RÓTULOS LIMITADOS

Luiz Fernando Souza Luz[1], *Matheus Pinheiro Ferreira*[2], *Laura Elena Cué la Rosa*[3], *Hudson Franklin Pessoa Veras*[4], *Ernandes Macedo da Cunha Neto*[5], *Evandro Orfanó Figueiredo*[6], *Ana Paula Dalla Corte*[2], *Carlos Roberto Sanquetta*[2] *e Dário Augusto Borges Oliveira*[1]

[1]School of Applied Mathematics, Getulio Vargas Foundation, Rio de Janeiro, RJ, Brazil, e-mail: luiz.fernando.luz608@gmail.com, [2]Department of Forest Sciences, "Luiz de Queiroz" College of Agriculture, University of São Paulo (USP/ESALQ), Piracicaba, SP, Brazil, [3]Laboratory of Geo-Information Science and Remote Sensing, Wageningen University and Research, Wageningen, The Netherlands; [4]Instituto Federal do Acre, Rio Branco, AC, Brazil; [5]Departament of Forest Science, Federal University of Paraná (UFPR), Curitiba, PR, Brazil; e [6]Embrapa Acre, Rio Branco, AC, Brazil

RESUMO

Este trabalho explora a aplicação de métodos de *self-sampling* usando modelos de Deep Learning no contexto de segmentação de instâncias de árvores por meio de imagens de sensoriamento remoto. O objetivo é desenvolver um método eficaz capaz de mapear diferentes espécies em áreas florestais com alta precisão, mesmo diante de um conjunto de dados inicial limitado. A abordagem proposta visa cobrir maiores áreas de floresta enquanto mantém a qualidade das instâncias segmentadas. A introdução dessas técnicas resultou em um aumento da cobertura de segmentação do modelo, de 4.4% para 23.7% da área de estudo, mantendo um F1-Score médio em patamares elevados, oscilando entre 60% e 45%.

Palavras-chave – *Self-Sampling*, *Deep Learning*, Sensoriamento Remoto, Monitoramento Florestal, Segmentação de Instância.

ABSTRACT

This work explores the application of a self-sampling method in Deep Learning models for instance segmentation of tree species using remote sensing images. The aim is to develop an effective method capable of mapping different tree species in forest areas with high precision, even when faced with a limited initial set of annotated data. The proposed approach aims to cover a larger area of the forest while maintaining satisfactory prediction quality. The introduction of these techniques resulted in an increase in the model's segmentation coverage from 4.4% to 23.7% of the study area, while maintaining a high macro-average F1-Score, ranging between 60% and 45%.

Key words – Self-Sampling, Deep Learning, *Remote Sensing, Forest Monitoring, Instance Segmentation.*

1. INTRODUÇÃO

A preservação das florestas é um tema de crescente preocupação global, devido ao aumento acelerado do desmatamento nessas áreas. Dentre os diversos tipos de desflorestamento, existe a degradação seletiva de espécies, uma atividade difícil de identificar em vegetação densa. Diante disso, faz-se necessário o mapeamento preciso de espécies para avaliar a presença desse tipo de prática, bem como seu impacto sobre a biodiversidade e o ecossistema local. Além disso, através do levantamento preciso da flora em áreas de recuperação, é possível monitorar o progresso das iniciativas de restauração e o restabelecimento da biodiversidade do ecossistema para sua composição natural.

Diante desse desafio, alguns estudos têm utilizado redes convolucionais para abordar desafios relacionados à identificação e segmentação de árvores. Em [1] utilizaram imagens de Veículos Aéreos Não Tripulados - VANT para esse tipo de mapeamento. Em [2] propuseram uma abordagem com *Fully Convolutional Networks* (FCN) para identificar a espécie *Cecropia hololeuca*, a partir de imagens de alta resolução. Em [3], utilizou-se FCN para classificar espécies de árvores em uma floresta de coníferas a partir de dados hiperespectrais, obtendo pontuações médias *F1-score* de 87%.

Porém, em sua grande maioria os modelos exigem abundância de dados anotados de copas de árvores (*Individual Tree Crown* - ITC) para obter resultados satisfatórios. Essa limitação inviabiliza a classificação de um vasto número de espécies em florestas densas, como as florestas tropicais, onde a proporção de espécies não é igualmente distribuída. Em geral, classes de árvores são subrepresentadas, resultando em poucas amostras de treino para as espécies mais raras [4], [5]. A coleta de amostras de espécies raras através de trabalho de campo, pode ser uma solução para essa problemática, porém essa alternativa exige tempo e demonstra-se bastante custosa, como mostra [6].

Visando gerar avanços na segmentação de ITC, o objetivo deste trabalho é introduzir técnicas de *self-sampling*. Nesse tipo de abordagem, a rede é inicialmente treinada em um conjunto reduzido de dados anotados. A partir disso, o modelo gera novas anotações que são adicionadas em sua base de treinamento e um novo ciclo de treino é iniciado. Nesse processo, aumenta-se gradualmente a quantidade de ITC identificadas, bem como a cobertura da área de estudo. Com esse método, é possível utilizar a eficácia dos modelos FCN a partir de um conjunto de dados limitado.

2. MATERIAL E MÉTODOS

2.1. Imagens de Sensoriamento Remoto

Nesse trabalho, foram utilizadas imagens coletadas por um veículo aéreo não tripulado (VANT), equipado com um sensor

Tabela 1: Summary of ITCs for different tree species.

Espécies	Nº ITC - Treino	Nº ITC - Teste
Abiorana Rosa	50	22
Angico	4	2
Angico Vermelho	9	5
Castanheira	58	25
Cedro	36	16
Cerejeira	9	5
Cumaru Ferro	14	6
Garapeira	31	14
Guaribeiro	16	8
Guariuba	12	6
Ipe	14	6
Jutai	22	10
Massaranduba	6	3
Samauma	7	4
Tachi	17	8
Tauari	46	21
Ucuuba	4	2

CMOS RGB de 20 megapixels e uma lente de 24 mm de foco automático e obturador manual. Realizou-se o levantamento a uma altitude de voo de 120 metros acima das copas da floresta, com o VANT mantendo uma velocidade de cruzeiro de 10 m/s sobre 260 hectares. Essa configuração resultou em imagens com *ground sampling distance* (GSD) de 4 cm, garantindo dados de alta resolução.

Antes dos voos, foram definidos três pontos de controle em solo ao redor da borda da reserva florestal, com o objetivo de assegurar a precisão geométrica. A coleta dos dados foi realizada por receptores GNSS de dupla frequência, que registraram informações dos GPS e GLONASS durante um período de 241 minutos. Após o processamento, os GCPs obtiveram uma precisão vertical média de 10 cm e uma precisão horizontal de 3 cm.

A coleta das imagens ocorreu exclusivamente na floresta experimental da Empresa Brasileira de Pesquisa Agropecuária (Embrapa), localizada no município de Rio Branco, no estado do Acre, Brasil (10º 01' 22"S, 67º 40' 3"O). Todos os voos de VANT foram realizados nesta área, que se caracteriza como uma floresta tropical aberta, com presença de árvores que excedem 30 metros de altura, bambus, diversas espécies de palmeiras e grandes árvores emergentes, apresentando uma distribuição típica de uma floresta tropical madura e não manejada.

2.2. Dados Rotulados com o Contorno das Copas de Árvores

Este trabalho fez uso de uma versão melhorada do banco de dados de ITC, criado por [7]. Esta versão abrange espécies de árvores com relevância econômica, tanto para a produção de madeira quanto de produtos não madeireiros. Os contornos das copas foram anotados a partir das imagens capturadas por VANT descritas em 2.1.. A identificação das espécies foi realizada com base em dados de inventário florestal e conhecimentos especializados em botânica. A Tabela 1 apresenta detalhes das espécies presentes no conjunto de dados.

2.3. Modelo Convolucional

Dentre os modelos conhecidos de segmentação semântica, optou-se pelo uso da rede DeepLabv3, proposta por [8]. A escolha dessa arquitetura foi embasada nos resultados obtidos por [9] em segmentação de copas de árvores em florestas densas. A versão escolhida DeepLabv3 possui um *encoder* com a ResNet50 [10] e dois *decoders* seguindo a estrutura proposta em [9], que visa lidar com o conjunto de dados limitados.

Adicionalmente, a disponibilidade dos pesos pré-treinados no conjunto *COCO Train 2017*, acessível através do repositório *PyTorch Hub*, constituiu um fator determinante na escolha da DeepLabv3. Durante os experimentos conduzidos neste estudo, a utilização desses pesos aprimorou o desempenho do modelo, particularmente diante da limitação de dados disponíveis para treinamento.

2.4. Processo Iterativo de Self-Sampling

O processo inicia-se com o conjunto de instâncias de árvores descrito em 2.2. Após passarem pelo pré-processamento descrito em [9], a rede convolucional é treinada.

Após o treinamento, o modelo conduz inferências em todo o conjunto de imagens sem anotação, resultando em um mapa de probabilidades da área. A partir dessa saída, são selecionadas as regiões nas quais o modelo apresenta confiança superior a 90% em sua classificação. Com essa seleção, obtêm-se novas ITC que são armazenadas em um inventário com todas as instâncias segmentadas.

Diferentemente do processo de *active learning*, onde as amostras adicionadas em cada iteração são selecionadas seguindo critérios específicos e rigorosos, no *self-sampling* a seleção é mais relaxada. As regras de seleção foram projetadas para eliminar apenas o ruído, sem impor restrições significativas à inclusão de novas instâncias.

Para a atualização do conjunto de treino, selecionam-se apenas 5 novas anotações por espécie do inventário. Essa seleção visa garantir o balanceamento dos dados de treino, assegurando a estabilidade do modelo. Durante esse acréscimo, assegura-se que as instâncias do treinamento inicial permaneçam presentes em todas as iterações, com o propósito de manter a influência dos dados originais sobre o modelo, evitando, assim, a geração de segmentações divergentes em relação às segmentações verdadeiras.

Ao final dessas etapas, o ciclo de *self-sampling* é reiniciado. Esse ciclo é realizado por 20 iterações, momento em que a quantidade de novas instâncias adicionadas ao inventário tornam-se irrelevantes.

2.5. Elaboração de Experimentos

Para realizar os experimentos, foram separadas 355 amostras para treino e 163 para teste, correspondendo, respectivamente, a 69% e 31% dos dados disponíveis.

No treinamento, foram selecionadas aleatoriamente janelas com dimensões de 512x512. Adotou-se a estratégia de *oversampling*, assegurando a uniformidade das espécies utilizadas no treino. Aplicaram-se técnicas de *augmentation* que incluem rotações aleatórias de 90°, 180° e 270°, bem como inversões verticais e horizontais. Além disso,

adicionou-se uma camada de *dropout* antes da camada de classificação, com uma taxa de 65%.

Em cada ciclo de *Self-Sampling*, o treinamento ocorreu em no máximo 30 *epochs*, contendo 2500 amostras por *epoch*, com a possibilidade de alterar o número de epochs devido ao uso de *early stopping*.

A interrupção de treinamento ocorreu quando a melhoria do *F1-score* no conjunto de validação era menor que 9e-5 por 5 epochs consecutivas. Para a otimização estocástica (SGD - *Stochastic Gradient Descent*), adotou-se um otimizador com *momentum* de 0.9 e uma taxa de aprendizado inicial de 0.1, com um decaimento de 0.1 a cada 5 *epochs*. No primeiro ciclo, a rede foi inicializada com pesos pré-treinados no conjunto *COCO Train 2017*.

Avaliou-se o desempenho com base no conjunto de teste, com amostras que não foram usadas pelo modelo durante a etapa de treinamento. O *Macro-Average F1-Score* foi adotado como principal métrica de avaliação dos resultados.

Os experimentos foram conduzidos no sistema operacional Ubuntu 22.04 LTS, utilizando uma CPU Intel Xeon Platinum 8176, 1000 GB de RAM e uma placa gráfica Tesla V100-PCIE-32GB. O modelo foi implementado utilizando o PyTorch *framework*.

3. RESULTADOS

3.1. Avaliação Quantitativa

Para uma análise comparativa, o modelo treinado apenas com os dados de treino inicial cobriu 4,4% da área de estudo, enquanto o modelo treinado com amostras de *self-sampling* compreendeu 23% da área (Figura 3). No que se refere à quantidade de ITC, o modelo evoluiu de 711 instâncias na primeira iteração, para 283 na 20ª iteração (Figura 4).

Ao longo dos ciclos, o F1-Score variou entre 60,6% e 44,2% (Figura 5). Apesar da queda observada no F1-Score, a métrica se manteve relativamente estável. Esse desempenho deve ser ponderado em relação ao aumento significativo, tanto no volume de ITC quanto na área segmentada, fato que indica a robustez do método.

3.2. Segmentação das Espécies

Entre os objetivos centrais deste estudo está o mapeamento da flora com eficácia, tornando importante a validação dos resultados obtidos pelo modelo nesse tópico.

A (Figura 6) evidencia que a espécie *Angico* não demonstrou aumento significativo na quantidade de amostras ao longo das iterações, mesmo com o emprego de técnicas de balanceamento de dados. Por outro lado, observa-se que espécies como *Massaranduba* e *Ucuuba*, que possuíam um conjunto inicial reduzido, mostraram um crescimento notável na quantidade de amostras. Esse resultado sugere que a proporção de amostras de cada espécie no conjunto inicial não é, necessariamente, um determinante para a distribuição das amostras ao término dos ciclos de *self-sampling*.

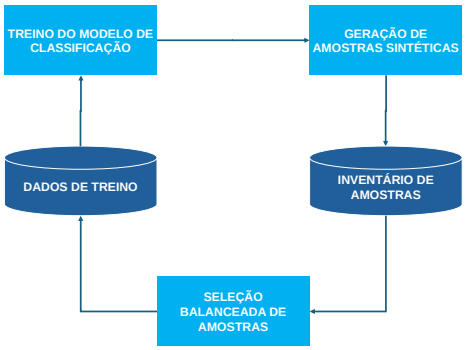

Figura 1: Etapas do método de self-sampling

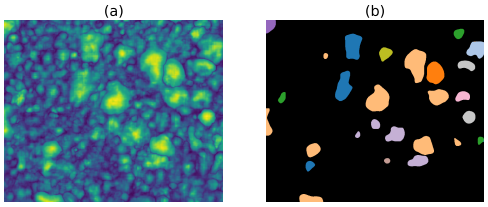

Figura 2: Exemplo de seleção de ITC. (a) Mapa de probabilidades gerado pelo modelo. (b) Instâncias selecionadas

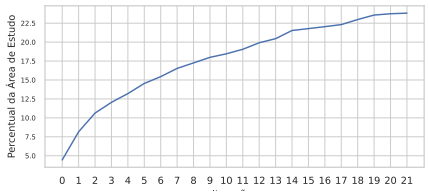

Figura 3: Evolução da área segmentada ao longo das iterações

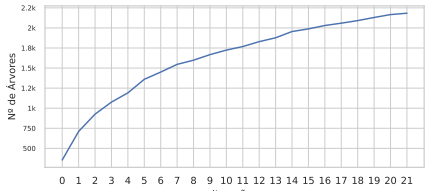

Figura 4: Evolução da quantidade de ITC segmentadas ao longo das iterações

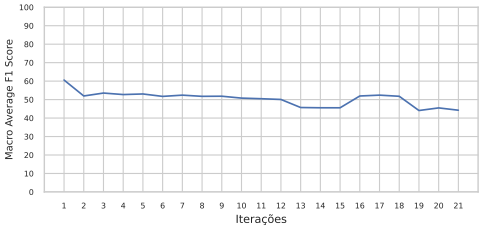

Figura 5: Evolução do F1-Score ao longo das iterações

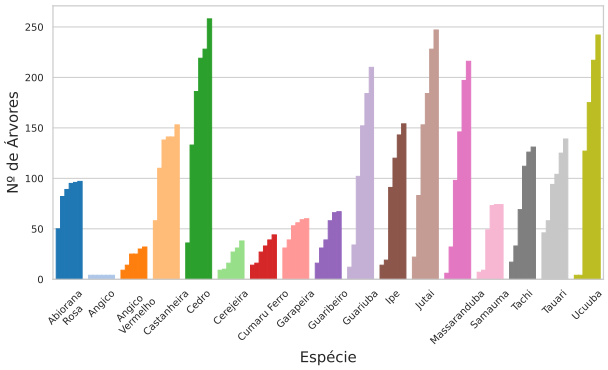

Figura 6: Evolução da Quantidade de ITCs por Espécie. Quantidades do conjunto de treino inicial, da 1ª, 5ª, 10ª, 15ª e 20ª iteração.

4. DISCUSSÃO

4.1. Potencial Uso em Escala Ampliada

Apesar de este estudo utilizar imagens de alta resolução, a aplicação de dados dessa qualidade, conforme discutido em 2.1, pode mostrar-se impraticável em escalas maiores, como em nível municipal ou estadual. Em cenários de vasta abrangência territorial, é comum recorrer a imagens de satélite com menor precisão e detalhamento, a fim de equilibrar a viabilidade técnica e os custos envolvidos na análise de grandes áreas.

Um passo relevante para a continuidade deste estudo é testar a viabilidade do método em escalas maiores. Caso o modelo apresente resultados satisfatórios, a segmentação de ITC poderá ser estendida a vastas áreas florestais, aproveitando a acessibilidade e a cobertura abrangente oferecidas pelos dados de satélite. Isso permitirá o monitoramento da flora em proporções significativamente maiores, sem comprometer a precisão das análises.

4.2. Monitoramento da Dinâmica Populacional da Flora

Além da aplicação em larga escala, essa ferramenta de segmentação tem potencial uso no monitoramento de espécies ao longo do tempo, proporcionando uma compreensão mais profunda da dinâmica populacional da flora.

Por meio de capturas recorrentes de imagens de uma mesma área, seria possível não apenas identificar a presença de espécies, mas também acompanhar variações em sua distribuição e densidade ao longo dos anos. Essa análise temporal forneceria métricas importantes sobre os padrões de crescimento, reprodução e possível declínio das espécies. Esse levantamento contínuo permitiria a identificação de espécies que estão enfrentando pressões ambientais, como mudanças climáticas, variações na disponibilidade de recursos ou impactos de atividades humanas, o que contribuiria para um entendimento abrangente do efeito desses fatores na vegetação.

A capacidade de monitorar a dinâmica populacional das espécies pode ser fundamental para a conservação da biodiversidade. Identificar quais espécies estão em declínio ou enfrentando pressões significativas permitiria a implementação de estratégias de manejo e conservação mais direcionadas e eficazes, visando à proteção dessas espécies vulneráveis.

5. CONCLUSÕES

Os resultados obtidos ao longo deste estudo demonstram a eficácia da aplicação de métodos de *self-sampling* combinados com técnicas de *deep learning* para a segmentação de ITC em imagens de sensoriamento remoto. Por meio da abordagem proposta, foi possível superar desafios associados à disponibilidade limitada de dados anotados, ampliando a área coberta pela segmentação e o número de instâncias identificadas, passando de 4,4% para 23% da área de estudo e de 139 instâncias para mais de 2.183 ITC. Ressalta-se que existem oportunidades de expansão e aprimoramento deste estudo. Investigações futuras poderiam explorar a aplicação dessas técnicas em diferentes biomas ou regiões geográficas, bem como a avaliação do desempenho do modelo em ambientes com maior diversidade e densidade de espécies arbóreas.

6. REFERÊNCIAS

[1] T. Kattenborn, J. Eichel, and F.E. Fassnacht. Convolutional neural networks enable efficient, accurate and fine-grained segmentation of plant species and communities from high-resolution uav imagery. *Scientific reports*, 9:1–9, 2019.

[2] F.H. Wagner, A. Sanchez, Y. Tarabalka, R.G. Lotte, M.P. Ferreira, M.P.M. Aidar, E. Gloor, O.L. Phillips, and L.E.O.C. Aragao. Using the u-net convolutional network to map forest types and disturbance in the atlantic rainforest with very high resolution images. *Remote Sensing in Ecology and Conservation*, 5:360–375, 2019.

[3] G.A. Fricker, J.D. Ventura, J.A. Wolf, M.P. North, F.W. Davis, and J. Franklin. A convolutional neural network classifier identifies tree species in mixed-conifer forest from hyperspectral imagery. *Remote Sensing*, 11:2326, 2019.

[4] A. Mellor, S. Boukir, A. Haywood, and S. Jones. Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin. *ISPRS Journal of Photogrammetry and Remote Sensing*, 105:155–168, 2015.

[5] Fabian Ewald Fassnacht, Hooman Latifi, Krzysztof Stereńczak, Aneta Modzelewska, Michael Lefsky, Lars T Waser, Christoph Straub, and Aniruddha Ghosh. Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, 186:64–87, 2016.

[6] S.J. Graves, G.P. Asner, R.E. Martin, C.B. Anderson, M.S. Colgan, L. Kalantari, and S.A. Bohlman. Tree species abundance predictions in a tropical agricultural landscape with a supervised classification model and imbalanced data. *Remote Sensing*, 8:161, 2016.

[7] Hudson Franklin Pessoa Veras, Matheus Pinheiro Ferreira, Ernandes Macedo da Cunha Neto, Evandro Orfanó Figueiredo, Ana Paula Dalla Corte, and Carlos Roberto Sanquetta. Fusing multi-season uas images with convolutional neural networks to map tree species in amazonian forests. *Ecological Informatics*, 71:101815, 2022.

[8] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. Rethinking atrous convolution for semantic image segmentation. *arXiv preprint arXiv:1706.05587*, 2017.

[9] Laura Elena Cué La Rosa, Camile Sothe, Raul Queiroz Feitosa, Cláudia Maria de Almeida, Marcos Benedito Schimalski, and Dario Augusto Borges Oliveira. Multi-task fully convolutional network for tree species mapping in dense forests using small training hyperspectral data. *arXiv:2106.00799v2 [cs.CV]*, Sep 2021.

[10] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.