

Um arcabouço de Seleção de Bandas Landsat-8 baseado em UMDA para Detecção de Desmatamento

Eduardo B. Neto, Paulo R. C. Pedro, Álvaro Fazenda, Fabio A. Faria
Universidade Federal de São Paulo – São José dos Campos, Brasil,
{ebneto, costa.paulo, alvaro.fazenda, ffaria}@unifesp.br

20 de novembro de 2023

Resumo

A conservação das florestas tropicais é um assunto atual de relevância social e ecológica, devido ao importante papel que elas desempenham no ecossistema global. Infelizmente, milhões de hectares são desmatados e degradados todo ano, sendo necessários programas (governamentais ou de iniciativas privadas) para monitoramento das florestas tropicais. Neste sentido, este trabalho propõe um novo arcabouço que utiliza de algoritmo de estimação de distribuição estocástico (UMDA) para selecionar bandas espectrais do satélite Landsat-8 que melhor representam as áreas de desmatamento que guiará a criação de novas composições de imagens para alimentar uma arquitetura de segmentação semântica chamada DeepLabv3+. Nos experimentos realizados foi possível encontrar diversas composições que conseguem acurácia balanceada maior que 90% na tarefa de classificação de segmentos. Além disso, a melhor composição de bandas (651) encontrada pelo algoritmo UMDA alimentou uma arquitetura DeepLabv3+ e resultou em melhor eficiência e eficácia contra todas as composições comparadas neste trabalho¹.

¹Códigos Fonte disponível em: <https://github.com/ebouhid/>.

1 Introdução

Florestas tropicais estão localizadas entre os Trópicos de Câncer e Capricórnio, próximas à linha do Equador, sendo encontradas na América do Sul e Central, na África e em regiões da Ásia e do Pacífico. Até 2015 existiam 700 milhões de hectares de floresta primária (original) [1]. Embora as florestas tropicais cubram apenas 7% da superfície terrestre, estima-se que elas abrigam mais da metade das espécies do planeta [1]. Além da grande biodiversidade, as plantas e o solo das florestas tropicais retêm de 460 a 575 bilhões de toneladas de carbono. Os processos de evaporação e evapotranspiração de suas plantas e árvores retornam grandes quantidades de água para a atmosfera local, promovendo a formação de nuvens e chuvas [2].

Infelizmente, milhões de hectares de florestas tropicais tem sido perdidos a cada ano através de desmatamento e degradação [1, 3]. De acordo com um dos mais conhecidos e bem sucedidos programas de monitoramento, PRODES (Programa de Monitoramento da Floresta Amazônica Brasileira por Satélite), no período entre agosto/2021 e julho/2022, o desmatamento na Amazônia Legal Brasileira alcançou 11.568 km^2 [4]. E para piorar o cenário, com a escassez de mão-de-obra especializada e abundância de dados a serem analisados, o processo de monitoramento da floresta se torna muito custoso tanto financeiramente quanto em relação ao tempo, o que constitui um grande desafio para as tecnologias da informação e comunicação (TIC).

Na literatura, técnicas de aprendizado de máquina e processamento digital de imagens têm sido desenvolvidas para superar os desafios encontrado e auxiliar os especialistas nas tarefas de detecção automática ou semi-automática de áreas de desmatamento em florestas tropicais. Ortega *et al.* [5], aponta que o uso de métodos de classificação baseados em aprendizado profundo – particularmente, Redes Neurais Convolucionais Siamesas e *Early Fusion* – é uma abordagem superior a técnicas tradicionais de Aprendizado de Máquina. Andrade *et al.* [6] aborda a detecção de desmatamento em imagens de satélite a partir da Segmentação Semântica (SS) utilizando o modelo DeepLabv3+. Ambos trabalhos utilizam imagens do satélite Landsat-8 com todas as sete bandas espectrais com adição de NDVI. Em Maretto *et al.* [7], foi proposta uma abordagem para detecção de desmatamento via segmentação semântica de imagens do Landsat-8 com fusão de imagens baseada na conhecida arquitetura U-Net. Para isso, foram utilizadas imagens compostas por 5 bandas (3, 4, 5, 6 e 7), escolhidas ad-hoc.

No sentido de explorar o real potencial de sensores multiespectrais tais como o Landsat-8, muitos trabalhos têm sido propostos para criar novas imagens de uma mesma área por meio da escolha de bandas espectrais e buscar as melhores representações para aplicação alvo. Yu *et al.* [8] propõe uma metodologia de seleção de bandas baseada em suas correlações para otimizar o desempenho de uma Máquina de Vetores de Suporte (SVM) na tarefa de classificação do uso/cobertura de terras. Já Dallaqua *et al.* [9], realizou um

estudo “força bruta” na busca da melhor composição de bandas que proporcionasse maior acurácia na detecção de desmatamento por um classificador SVM. Nesse contexto, o uso de algoritmos de busca se mostra como uma boa solução alternativa, por ser mais eficiente em tempo e custo computacional do que métodos tradicionais, como *grid search* convencional [10]. Dentre os trabalhos existentes na literatura que abordam esse tema destaca-se Nagasubramanian *et al.* [11] que faz uso de algoritmos genéticos na busca da melhor combinação de bandas de um sensor hiperespectral para a detecção de pragas em plantações de soja.

Nesse sentido, este trabalho propõe um arcabouço de seleção de bandas espectrais baseado em UMDA (*Univariate Marginal Distribution Algorithm*) para servir como guia para a construção da base de imagens de entrada para arquiteturas de segmentação semântica de imagens de desmatamento da Amazônia.

2 Um arcabouço de Seleção de Bandas LandSat-8 baseado em UMDA para Detecção de Desmatamento

A Figura 1 mostra em detalhes o funcionamento do arcabouço proposto neste trabalho. Em (a) está uma imagem do satélite Landsat-8 com suas sete bandas espectrais, (b) é a imagem falsa-cor formada de três componentes principais do PCA [12], (c) é a imagem resultante da segmentação SLIC [13] com diferentes segmentos encontrados, (d) é a imagem resultante da seleção de segmentos indicados pelos pontos brancos, (e) estão os vetores de características Haralick [14] extraídos dos segmentos selecionados, (f) estão os vetores de características separados em três conjuntos (treino, validação e teste) para serem utilizados no treinamento do classificador SVM [15] durante processo evolutivo do algoritmo UMDA [16], (g) está o melhor indivíduo encontrado pelo algoritmo UMDA (bandas 4, 6 e 7), (h) está a composição da imagem representada pelo melhor indivíduo e finalmente em (i) está a imagem segmentada pela arquitetura DeepLabv3+ [17], onde os *pixels* brancos correspondem às regiões de desmatamento na imagem de entrada.

2.1 Criação da Base de Segmentos

A base de dados criada para este trabalho é composta por 10 diferentes imagens do satélite Landsat-8 da região da Bacia do Xingu que totaliza mais de 9.497 km^2 de floresta. Para a realização dos experimentos foi adotado o mesmo protocolo experimental de [9]. Portanto, cada uma das 10 imagens composta por 7 bandas espectrais foram submetidas ao algoritmo de redução de dimensionalidade PCA passando de 7 para apenas 3 bandas (3 primeiras

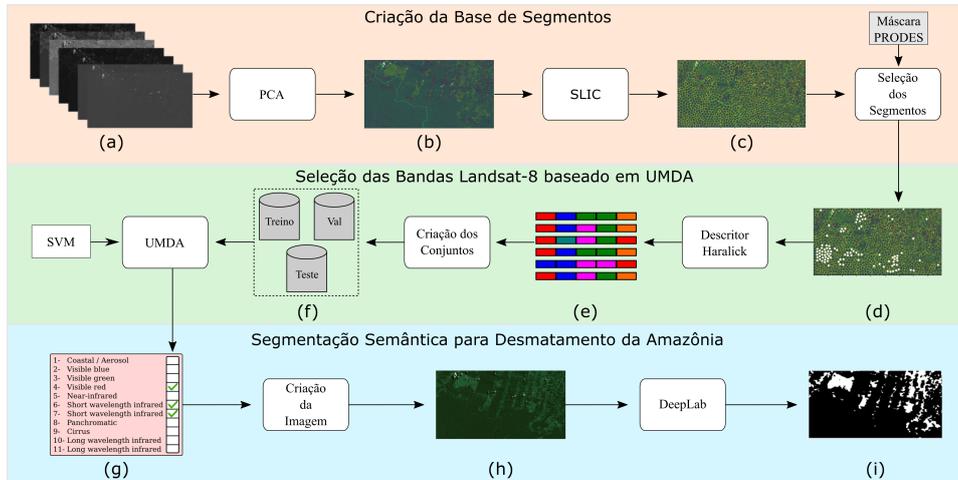


Figura 1: O arcabouço de seleção de bandas Landsat-8 baseado em UMDA para detecção de desmatamento.

componentes principais) e um algoritmo de segmentação não-supervisionada SLIC [13] foi aplicado para criação de 11831 segmentos (11404 de “Floresta” e 427 de “Não-floresta”). Uma vez com segmentos criados, realizou-se uma etapa de seleção dos melhores segmentos com adoção de dois critérios: (1) segmentos com Taxa de Homogeneidade (proporção da classe majoritária) $\geq 0,70$ considerando a verdade fornecida pelo PRODES [4]; e (2) segmentos com área mínima de 70 pixels . A Tabela 1 descreve as 10 regiões com diferentes dimensões e suas representações em km^2 . Os valores de Floresta e Não-floresta estão relacionados com a quantidade de segmentos encontrados pelo SLIC para cada uma das classes.

2.2 Seleção das Bandas Landsat-8 baseada em UMDA

Para cada um dos 11831 segmentos, são calculados seus respectivos descritores de textura Haralick [14] compostos de 13 coeficientes calculados em 4 direções para cada uma das 7 bandas espectrais, o que totaliza 7 vetores de características para cada segmento. Então, os segmentos são separados em três conjuntos (70% de treino, 15% de validação e 15% de teste), os quais serão utilizados no processo evolutivo do algoritmo UMDA. Introduzido por Mühlenbein & Paaß [16], o algoritmo *Univariate Marginal Distribution* (UMDA) é um dos mais simples algoritmos de estimação de distribuição (AED) que assumem independência entre variáveis de decisão e tem obtido resultados superiores aos alcançados pelos algoritmos genéticos tradicionais [18]. Em seu processo evolutivo, cada indivíduo da população é composto de 7 genes que representam as bandas espectrais que compõem uma imagem Landsat-8. A presença da banda no indivíduo é marcada com valor 1 e sua ausência com 0. O algoritmo inicia com uma população aleatória,

Tabela 1: Descrição detalhada das imagens Landsat-8 coletadas para este trabalho.

Região	km²	Dimensões	Floresta	Não-floresta	Total
1	930,40	1230× 843	1585	107	1692
2	1124,59	1343× 933	1115	134	1249
3	875,36	977× 998	1167	25	1192
4	607,33	768× 879	475	9	484
5	982,86	790× 1384	612	36	648
6	694,77	928× 833	1104	17	1121
7	1528,65	1788× 950	2016	9	2025
8	778,84	853× 1017	928	36	964
9	925,91	971× 1064	1582	19	1601
10	1048,45	1047× 1115	820	35	855
Total	~9.497	–	11404	427	11831

seguido de cálculo da função de aptidão de cada indivíduo. A função de aptidão utilizada é a acurácia balanceada de um classificador SVM (*Support Vector Machine*), empregando segmentos formados pelos descritores de Haralick referentes às bandas do indivíduo em questão. Em seguida, os n melhores indivíduos são selecionados para compor o conjunto de indivíduos “pais”. As probabilidades marginais são calculadas a partir da frequência relativa de cada gene presente entre todos os pais, estimando sua distribuição de probabilidade. Considerando essa distribuição de cada gene, novos indivíduos são gerados como “prole”. A população da próxima geração é composta pela união dos conjuntos de indivíduos da “prole” e dos “pais”. O processo é repetido até que o critério de parada seja atingido, sendo a quantidade de gerações o critério adotado neste estudo. Dentre as diversas configurações testadas para o algoritmo UMDA visando eficiência no processo evolutivo, os melhores parâmetros encontrados foram: 10 indivíduos na população, 10 gerações, número de indivíduos selecionados (pais) igual a 5 e número descendentes gerados igual a 5.

2.3 Segmentação Semântica para Desmatamento da Amazônia

Nesta etapa, as 10 diferentes imagens originais presentes na base de dados Landsat-8, descritas na Tabela 1, são reconstruídas por meio da composição das bandas mais relevantes (isto é, aquelas com maior representatividade dentre os melhores indivíduos do UMDA, conforme a Tabela 2). Por sua vez, essas imagens são separadas em dois conjuntos, sendo que o treino é composto de 8 regiões (1, 2, 5, 6, 7, 8, 9 e 10) e o teste composto por 2 regiões (3 e 4). Similar ao trabalho [6], essa configuração dos conjuntos de treino e

teste se deram de forma manual, uma vez que é desejável que imagens com nuvens não estivessem presentes no conjunto de teste.

Durante processo de treinamento da arquitetura de segmentação semântica DeepLabv3+ [17], cada imagem dos conjuntos de treino e teste é recortada em sub-imagens de dimensões 256×256 com *stride* de 64 *pixels*. O recurso de *data augmentation* foi utilizado para o conjunto de treino, que sofreu transformações de rotação, apenas. Com isso, obteve-se 10728 imagens para treino e 234 para teste.

3 Resultados & Discussão

Esta seção discute em detalhes os resultados dos experimentos realizados neste trabalho.

3.1 Classificação de Segmentos pelo Classificador SVM

Neste experimento, realizou-se uma análise comparativa dos resultados de classificação do SVM utilizando os melhores indivíduos encontrados pelo algoritmo UMDA.

A Tabela 2 mostra apenas os indivíduos encontrados pelo algoritmo UMDA com acurácia balanceada maior que 90,0% obtida pelo classificador SVM no conjunto de teste. É possível observar, como em todo algoritmo não-determinístico, os resultados de acurácia dos indivíduos/soluções variam dependendo da semente escolhida no experimento. Entretanto, um fato muito interessante de comentar é a presença frequente das bandas 6, 5 e 1, observadas 92,3%, 92,3% e 69,2% dos indivíduos, respectivamente. Portanto, pode-se sugerir a grande importância dessas bandas para a tarefa alvo.

A Tabela 3 mostra uma análise comparativa entre os melhores indivíduos encontrados pelo algoritmo UMDA e três composições de bandas comumente utilizadas na literatura com as cinco diferentes sementes de inicialização. É possível notar que os resultados do classificador SVM utilizando os melhores indivíduos conseguem resultados similares aos da melhor composição da literatura, a qual utiliza de todas as sete bandas espectrais. Já quando comparados os melhores indivíduos com as composições RGB e PCA, os ganhos relativos são 11,88% e 5,93%, respectivamente.

3.2 Segmentação Semântica de Imagens de Desmatamento da Amazônia

Visando enfatizar a importância das bandas espectrais mais frequentes encontradas no experimento anterior para a tarefa de classificação de segmentos de desmatamento na Amazônia, a Tabela 4 apresenta uma análise comparativa da segmentação semântica usando a arquitetura DeepLabv3+

Tabela 2: Resultados do classificador SVM com os melhores indivíduos encontrados pelo algoritmo UMDA. O símbolo ✓ significa a presença da banda espectral no indivíduo.

Semente	Indivíduo UMDA	Bandas							Acurácia Balanceada
		1	2	3	4	5	6	7	
10	1	✓		✓	✓	✓	✓		93,74
42	2		✓		✓	✓	✓		93,24
30	3	✓		✓	✓	✓	✓	✓	93,12
30	4	✓		✓		✓	✓	✓	92,71
20	5		✓	✓	✓	✓	✓	✓	92,67
20	6	✓		✓	✓	✓	✓	✓	92,58
42	7			✓	✓	✓	✓		92,51
10	8	✓				✓	✓	✓	92,35
42	9	✓				✓		✓	91,63
1	10	✓		✓	✓	✓	✓		91,50
1	11				✓	✓	✓	✓	91,39
30	12	✓		✓			✓	✓	90,69
20	13	✓	✓			✓	✓		90,22
Prob.	-	69,2%	23,1%	61,5%	61,5%	92,3%	92,3%	61,5%	-
Ordem	-	3°	7°	4°	5°	2°	1°	6°	-

com diferentes variações das bandas mais frequentes. Em todos os experimentos, utilizamos como base a arquitetura DeepLabv3+ com os pesos do *encoder* pré-treinados na base de dados ImageNet [19]. Em seguida, realizamos o ajuste fino do modelo com os seguintes parâmetros: 100 épocas de treinamento, taxa de aprendizado de 5×10^{-4} e função de custo "Jaccard Loss". Além disso, outras três composições da literatura foram adicionadas à análise (todas+NDVI [20], PCA e RGB). Como pode ser observado, o melhor resultado na tarefa alvo foi obtido pela composição das três bandas espectrais mais frequentes do estudo anterior (6, 5 e 1), conseguindo superar todas as três composições da literatura comparadas neste experimento, inclusive a composição de todas as 7 bandas do Landsat-8 + NDVI proposta por [20] para 3 de 4 medidas de avaliação calculadas e, também, no tempo de convergência da arquitetura DeepLabv3+. Este fato sugere que a seleção de bandas espectrais baseada em UMDA pode trazer grandes benefícios de eficiência (ganho de tempo) e eficácia (qualidade de resultados) para a tarefa de segmentação semântica de imagens de desmatamento da Amazônia.

4 Conclusões

Neste trabalho, um novo arcabouço tem sido proposto para detecção de áreas de desmatamento da Amazônia. Por meio da seleção de bandas do satélite Landsat-8 baseada no algoritmo estocástico UMDA, foi possível encontrar as bandas que mais contribuem para a representação das áreas de interesse.

Tabela 3: Resultado dos melhores indivíduos para cada semente e três composições conhecidas da literatura.

Semente	Indivíduo UMDA	Todas (7 Bandas)	RGB	PCA [9]
1	91,50	89,42	82,21	85,41
10	93,74	91,76	84,44	88,47
20	92,67	93,63	80,23	86,30
30	93,12	94,17	84,90	88,94
42	93,24	94,83	83,82	89,84
Média	92,85	92,76	83,12	87,79
Desvio Padrão	0,85	2,19	1,91	1,86

Tabela 4: Resultados da segmentação semântica utilizando DeepLab.

Composição	Precision	Recall	F1-score	Acurácia	mIoU	Época
6	88,00	89,68	88,83	87,68	79,91	23
6 e 5	89,21	89,74	89,47	88,46	80,95	5
6, 5 e 1	86,76	93,22	89,87	88,52	81,61	5
6, 5, 1 e 3	90,41	88,30	89,34	88,49	80,74	13
6, 5, 1 e 4	86,83	91,29	89,00	87,68	80,19	2
6, 5, 1 e 7	89,29	90,16	89,72	88,71	81,36	10
Todas + NDVI [20]	87,92	90,46	89,17	87,99	80,45	61
RGB	87,41	91,05	89,19	87,94	80,49	39
PCA	91,01	87,00	88,96	88,20	80,12	5

Uma vez com esse conhecimento, as imagens da base foram reconstruídas utilizando apenas as bandas espectrais mais relevantes e essas imagens alimentaram uma arquitetura de segmentação semântica (DeepLabv3+) para tarefa de detecção de desmatamento. Nos experimentos realizados, foi possível mostrar que o arcabouço de seleção de bandas espectrais proposto pode ser uma solução para conseguir maior eficiência (redução de tempo de treinamento) e eficácia (qualidade de resultados) quando comparado com abordagens comumente utilizadas na literatura para a tarefa alvo.

Agradecimentos

Os autores gostariam de agradecer a equipe do Projeto *ForestEyes* [9] pela colaboração, discussão e criação da base de dados de desmatamento. Além disso, agradecer a agência de fomento CNPq pela bolsa PIBIC do aluno Eduardo B. Neto. Esta pesquisa é parte do INCT of the Future Internet for Smart Cities financiada pelo CNPq (processo #465446/2014-0), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001 e FAPESP (processos #2018/23908-1,

#2019/26702-8, #2014/50937-1 e #2015/24485-9). Esse trabalho usou recursos do “SDumont - Sistema de Computação Petafópica do SINAPAD”.

Referências

- [1] C. Martin, *On the Edge: The State and Fate of the World’s Tropical Rainforests*. Greystone Books Ltd, 2015.
- [2] G. Urquhart, W. Chomentowski, D. Skole, and C. Barber, “Tropical deforestation,” 1998.
- [3] M. Hansen et al., “High-resolution global maps of 21st-century forest cover change,” *science*, vol. 342, no. 6160, pp. 850–853, 2013.
- [4] INPE, “Estimativa de desmatamento na Amazônia Legal para 2022 é de 11.568 km^2 ,” <https://encurtador.com.br/aBDP2>, 2022, accessed: 2023-08-01.
- [5] M. Ortega, J. Bermudez, P. Happ, A. Gomes, and R. Feitosa, “Evaluation of deep learning techniques for deforestation detection in the amazon forest,” *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 4, pp. 121–128, 2019.
- [6] R. B. Andrade et al., “Evaluation of semantic segmentation methods for deforestation detection in the amazon,” *ISPRS Archives; 43, B3*, vol. 43, no. B3, pp. 1497–1505, 2020.
- [7] R. V. Maretto, L. M. Fonseca, N. Jacobs, T. S. Körting, H. N. Bendini, and L. L. Parente, “Spatio-temporal deep learning approach to map deforestation in amazon rainforest,” *IEEE Geoscience and Remote Sensing Letters*, vol. 18, no. 5, pp. 771–775, 2020.
- [8] Z. Yu, L. Di, R. Yang, J. Tang, L. Lin, C. Zhang, M. S. Rahman, H. Zhao, J. Gaigalas, E. G. Yu *et al.*, “Selection of landsat 8 oli band combinations for land use and land cover classification,” in *2019 8th International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*. IEEE, 2019, pp. 1–5.
- [9] F. Dallaqua, A. Fazenda, and F. Faria, “ForestEyes Project: Can Citizen Scientists Help Rainforests?” in *IEEE 15th International Conference on eScience*. IEEE, 9 2019, pp. 18–27.
- [10] J.-P. Ma, Z.-B. Zheng, Q.-X. Tong, and L.-F. Zheng, “An application of genetic algorithms on band selection for hyperspectral image classification,” in *Proceedings of the 2003 international conference on machine learning and cybernetics (IEEE Cat. No. 03EX693)*, vol. 5. IEEE, 2003, pp. 2810–2813.

- [11] K. Nagasubramanian, S. Jones, S. Sarkar, A. K. Singh, A. Singh, and B. Ganapathysubramanian, “Hyperspectral band selection using genetic algorithm and support vector machines for early identification of charcoal rot disease in soybean stems,” *Plant methods*, vol. 14, pp. 1–13, 2018.
- [12] I. Jolliffe, *Principal component analysis*. Springer, 2011.
- [13] R. Achanta et al., “SLIC superpixels compared to state-of-the-art superpixel methods,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.
- [14] R. M. Haralick et al., “Textural features for image classification,” *IEEE Transactions on systems, man, and cybernetics*, no. 6, pp. 610–621, 1973.
- [15] B. E. Boser et al., “A training algorithm for optimal margin classifiers,” in *Workshop on Computational Learning Theory*, ser. COLT ’92, 1992, pp. 144–152.
- [16] H. Mühlenbein and G. Paass, “From recombination of genes to the estimation of distributions i. binary parameters,” in *International conference on parallel problem solving from nature*. Springer, 1996, pp. 178–187.
- [17] Liang-Chieh Chen et al., “Encoder-decoder with atrous separable convolution for semantic image segmentation,” 2018.
- [18] Á. R. Ferreira, G. H. de Rosa, J. P. Papa, G. Carneiro, and F. A. Faria, “Creating classifier ensembles through meta-heuristic algorithms for aerial scene classification,” in *2020 25th International Conference on Pattern Recognition (ICPR)*, 2021, pp. 415–422.
- [19] J. Deng et al., “Imagenet: A large-scale hierarchical image database,” in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.
- [20] D. Torres et al., “Deforestation detection with fully convolutional networks in the amazon forest from landsat-8 and sentinel-2 images,” *Remote Sensing*, vol. 13, no. 24, 2021.