

Classificação Semiautomática de Áreas Queimadas com o uso de Redes Neurais

Ronaldo Nelis de Andrade, Olga Bittencourt, Fabiano Morelli, Rafael Santos

Instituto Nacional de Pesquisas Espaciais (INPE)
Av. dos Astronautas, 1758 – 12.227-010 – São José dos Campos – SP – Brazil
rondweb@gmail.com,
{olga.bittencourt,fabiano.morelli,rafael.santos}@inpe.br

***Abstract.** This paper presents an approach to improve the semi-automatic detection of burned areas through the use of neural networks. The approach is validated over a selected study area in the Brazilian Cerrado against reference data derived from data classified by experts. Methods are still being developed and improved, and initial results corroborate the validity of the approach, which will be extended to other study areas.*

***Resumo.** Esse trabalho apresenta uma abordagem para melhorar a classificação semiautomática de áreas queimadas através do uso de redes neurais. A validação foi realizada com um estudo de caso de uma área do Cerrado Brasileiro e os resultados foram comparados com áreas classificadas manualmente por especialistas. A pesquisa continua sendo aprimorada e os resultados iniciais corroboram o emprego da abordagem, que será ampliada para outras área de estudo.*

1. Introdução

O Cerrado é a região de savana rica em biodiversidade e um dos biomas mais ameaçados do país. Sua área ocupa em torno 204 milhões de hectares, 24% do território brasileiro, e já perdeu quase metade de sua cobertura vegetal original. Desmatamentos e queimadas são os maiores responsáveis por esse processo, tendo sido registrados mais de 30.000 focos de incêndio por ano neste bioma nos últimos 15 anos. Analisar os aspectos relacionados à ocorrência do fogo e aos impactos econômicos, sociais e ambientais gerados é um problema relevante e tem motivado pesquisas em áreas distintas do planeta. [Bowman et al, 2009; Katagis et al, 2014].

Nos últimos anos, foi desenvolvida uma nova geração de satélites capazes de fornecer imagens com melhores resoluções e georeferenciamento mais preciso, como CBERS-4 Landsat 8 e Sentinel-2. Li & Roy apresentam características e mostram os avanços obtidos por alguns satélites dessa nova geração [Li & Roy, 2017]. Avanços na utilização de imagens de sensoriamento remoto e na forma como o monitoramento de queimadas é realizado podem ser vistos em [Melchiori et al, 2014; Boschetti et al, 2015].

Nesse contexto de surgimento de novas tecnologias que permitem monitorar e quantificar mais precisamente as queimadas, o Instituto Nacional de Pesquisas Espaciais (INPE) desenvolve um programa de Monitoramento de Queimadas e Incêndios Florestais [INPE, 2017] que analisa continuamente aspectos relacionados à ocorrência

de fogo em áreas de vegetação tanto no bioma Cerrado quanto no restante do território brasileiro e parte da América Latina. São disponibilizados diversos produtos de dados, como o monitoramento diário de focos de incêndio, previsão de risco de fogo e estimativas periódicas de emissão de poluentes e, mais recentemente, estimativas de superfícies queimadas. Seus resultados são utilizados, por exemplo, como subsídios de políticas públicas como o Código Florestal e para contribuir para que as metas de redução das emissões de gases assumidas pelo Governo brasileiro na Convenção do Clima [MPOG,2015] possam ser atingidas.

Para produzir resultados cada vez mais precisos é importante identificar melhor as localizações, extensões e gravidades das queimadas. Seguindo essa linha de pesquisa, apresentamos aqui um trabalho em andamento que possui como contribuição inicial a diminuição do número de falsas detecções de áreas queimadas utilizando uma abordagem baseada no uso de redes neurais. A validação foi realizada com um estudo de caso de uma área do Cerrado Brasileiro e os resultados foram comparados a áreas classificadas manualmente por especialistas. A abordagem continua sendo refinada e trabalhos futuros de melhorias são discutidos.

2. Detecção de áreas queimadas usando Imagens Orbitais de Média Resolução

Entre os satélites de observação da Terra que disponibilizam imagens orbitais de média resolução destaca-se o Programa Landsat, que, desde a década de 80, gera imagens da mesma área a cada 16 dias. Seu mais novo satélite, o Landsat 8, disponibiliza imagens com resolução espacial na faixa de 30m para cada pixel. As imagens podem adquiridas na base de dados do Serviço Geológico dos Estados Unidos [USGS, 2016].

O desenvolvimento de abordagens que utilizam imagens de sensoriamento remoto para detectar mudanças ocasionadas por fogo apresenta muitos desafios. Em locais com grande extensão territorial ou áreas de difícil acesso, como o Brasil, a utilização dessas imagens e suas propriedades espectrais são a forma mais eficiente para monitorar queimadas. [Pereira et al, 2016] apresenta uma comparação da utilização de índices espectrais na detecção de áreas queimadas e ressalta fatores que interferem na qualidade da detecção, como as diferenças entre biomas e dentro de cada bioma e também o tempo de ocorrência entre o incêndio e a aquisição da imagem. Alguns trabalhos para estimar áreas queimadas foram desenvolvidos para analisar eventos específicos [Katagis et al, 2014]. Outros autores propõem minimizar as diferenças incorporando características regionais nas análises [Libonati et al, 2015; Boschetti et al, 2015].

2.1. O monitoramento de queimadas realizado pelo INPE

O INPE utiliza um método semiautomático para detectar áreas que sofreram mudanças de cobertura de vegetação [Melchiori et al, 2014]. O processo começa com a avaliação de cada pixel da imagem e seus índices de vegetação NDVI (Índice de Vegetação Normalizada) e NBRL (Índice de Queimada Normalizada). Se esses valores estiverem dentro de limites estabelecidos por especialistas, eles indicam que os pixels pertencem a possíveis queimadas. Após esse passo, realiza-se a segmentação através da verificação dos pixels vizinhos mais semelhantes aos pixels indicados como queimadas com a posterior delimitação de área de abrangência daquela possível queimada.

Essa abordagem apresenta como limitação o fato de ser semiautomática, ou seja, é necessária a revisão dos resultados por um especialista antes da sua publicação oficial. Essa verificação tem sido realizada manualmente e, apesar de apresentar bons resultados, é um processo dispendioso pois cada cicatriz deve ser reavaliada. Além disso, ao estendermos essa metodologia para todo o bioma Cerrado, o método apresenta um número de falsas detecções maior que o desejado em torno de 10%.

Uma hipótese para as confusões geradas pela metodologia de Melchiori é o fato de ela ter sido inicialmente produzida para atender áreas de preservação, onde todas as mudanças na vegetação, teoricamente, são decorrentes do uso do fogo. Assim, quando foi ampliada a área de estudo, novos elementos de mudança da paisagem, especialmente ligadas com atividades humanas passaram a ser monitoradas causando maior dificuldade para a acurácia da metodologia.

A Figura 1 ilustra áreas corretamente indicadas como queimadas na cor vermelha e, na cor azul, áreas de confusão. Essas últimas, muitas vezes indicam mudanças que aconteceram na vegetação, mas não foram originadas por fogo, por exemplo, colheitas em áreas agrícolas.

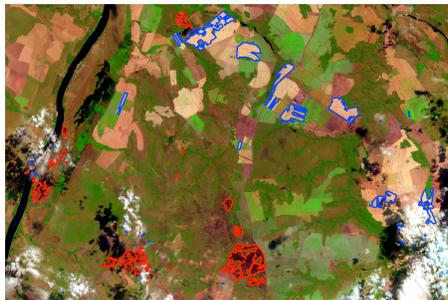


Figura 1. Fragmento contendo indicações de áreas queimadas (destacadas em vermelho) e falsas detecções (destacadas em azul).

Para contribuir com a diminuição do número de falsas detecções este trabalho apresenta uma abordagem que reavalia as cicatrizes e automaticamente separa as cicatrizes de queimadas, das cicatrizes de áreas de colheita e desmatamento, por exemplo. Serão apresentados os resultados preliminares alcançados com a utilização de redes neurais para filtrar os dados gerados pelo emprego do algoritmo desenvolvido por Melchiori [Melchiori et al, 2014].

3. Utilização de Redes Neurais na classificação de áreas queimadas

Redes Neurais podem ser definidas como técnicas computacionais que apresentam um modelo inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência. São modeladas através de dois processos: treinamento e aprendizagem. O treinamento é um processo iterativo de ajustes aplicados a seus pesos. O aprendizado é quando a rede neural atinge uma solução generalizada para uma classe de problemas.

A abordagem se baseia na utilização de redes neurais para extrair conhecimento de um conjunto de dados previamente validado, composto por cicatrizes de queimadas e falsas detecções. Essas redes são capazes de adquirir o conhecimento e gerar um modelo de classificação de queimadas. O objetivo é utilizar a base histórica de conhecimento de

áreas queimadas para auxiliar a verificação do resultado de novas análises. Neste estudo usamos o pacote NNET [Venables, W. N. & Ripley, B. D. (2002)] da linguagem R para implementar uma rede neural baseada em *perceptron* de múltiplas camadas com uma camada escondida.

Para comprovar a eficácia da abordagem escolhemos 6 imagens da órbita-ponto 222_066, parte do bioma Cerrado no estado do Tocantins. Cada imagem representa no solo uma área de abrangência de 18.500 x 18.500 ha. Desse conjunto, 3 imagens correspondem ao período de Julho a Setembro de 2015 e as outras 3, correspondem ao período de Julho a Setembro de 2016.

Utilizamos um conjunto de 3 meses para incorporar semelhanças climáticas e do solo do mesmo período do ano, considerando as alterações sazonais no terreno como: exposição do solo, umidade e ações antrópicas recorrentes. Testamos atributos espectrais relacionados às cicatrizes e escolhemos os 4 atributos mais representativos para compor a rede neural: mediana das bandas 5 e 6 e os índices NDVI e NBRL.

4. Resultados preliminares

No resultado do algoritmo original de detecção de queimadas, antes da validação manual, foram indicadas 10.299 áreas queimadas em 2015, correspondendo a uma área de aproximadamente 104.944 ha. Desse conjunto, a classificação manual identificou 5.935 classificações corretas e 4.364 falsas detecções. No ano de 2016, foram identificados inicialmente 9.029 polígonos que correspondem a 186.608 ha. Desse conjunto, a validação manual identificou 3.714 classificações e 5.315 falsas indicações. A Figura 2 exhibe detalhes das indicações de cicatrizes em total de polígonos e os totais de áreas. Notamos aqui que o número de polígonos de falsas detecções é alto, apesar de representarem uma pequena área em relação à área total analisada, indicando que a maior parte das falsas detecções ocorre nas menores cicatrizes.

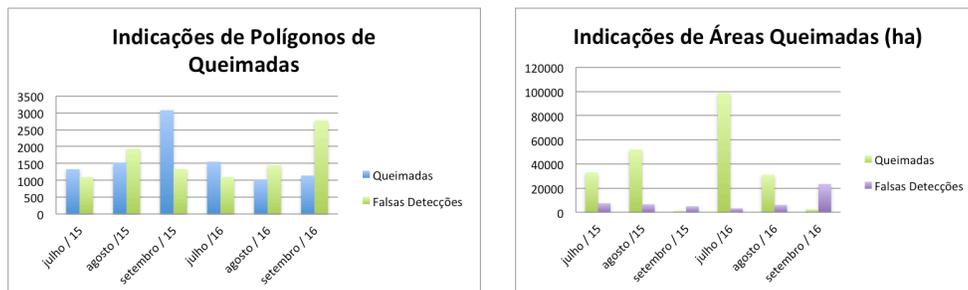


Figura 2. Gráficos mostrando os valores mensais de queimadas e falsas detecções em relação à quantidade de polígonos e à área analisada.

Os dados do ano de 2015 foram utilizados para construir a rede de conhecimento e gerar um modelo de classificação. Utilizamos dois tipos de classes: queimada e falsa detecção. Os dados de 2016, sem a classificação final, foram utilizados como entrada do modelo de redes neurais, para que fossem classificados pela rede como queimadas ou falsas detecções. Comparamos os resultados gerados pelo classificador com a validação manual. A Tabela 1 mostra os resultados considerando os índices de produtor (dos dados pertencentes a cada classe, quantos foram corretamente classificados), usuário (dos dados que foram classificados pelo modelo, quantos estavam certos dentro de cada classe) e acurácia global (desempenho geral da classificação).

Tabela 1. Resumo dos resultados da classificação usando Redes Neurais

Meses	Queimada			Falsa Detecção			Acurácia Global (%)
	Cicatrizes	Usuário (%)	Produtor (%)	Cicatrizes	Usuário (%)	Produtor (%)	
Jul	1.325	96,40	83,00	2.590	91,82	98,42	93,15
Ago	837	77,07	94,38	1.635	96,75	85,63	88,59
Set	1.731	94,77	84,86	911	76,00	91,11	87,00

Os resultados preliminares mostram acurácias próximas ao valor desejado de 90%. Notamos que para conjuntos menores de amostras, como as queimadas em agosto e falsas detecções em setembro, a precisão do resultado é menor. Estamos verificando a hipótese de ser necessário um conjunto mínimo de dados para realizar o treinamento.

A Tabela 2 mostra os resultados da matriz de confusão. Notamos que do total de 9029 cicatrizes indicadas, o modelo foi capaz de comprovar 3357 queimadas e encontrar 4779 falsas detecções. Essa matriz nos permite calcular métricas que indicam uma visão geral do desempenho da abordagem de aprendizado.

Nesse experimento, a sensibilidade é de 0.86 e mostra a capacidade do modelo em prever corretamente a condição de ser uma queimada para todos os polígonos que realmente são queimadas. A especificidade é de 0.93 e mostra a capacidade em prever corretamente as falsas detecções para os casos que realmente não são queimadas. A acurácia é de 0.90 e mostra uma proporção geral das predições corretas do modelo. Avaliamos que resultados em torno de 0.9 atendem nossa expectativa e consideramos o modelo gerado pelas redes neurais capaz de treinar seu conjunto de dados e distinguir adequadamente queimadas e falsas detecções.

Tabela 2. Matriz de confusão dos resultados totais obtidos

		Indicados pelo especialista	
		Queimada	Falsa Detecção
Previstos pelo modelo	Queimada	3357	357
	Falsa Detecção	536	4779

Para diminuirmos os erros de classificação estamos testando outras configurações para as redes neurais, como a inclusão de novos neurônios, que possam melhorar os resultados e não aumentar o tempo de processamento. Também estamos analisando outros atributos dos dados que possam enriquecer o modelo gerado.

5. Conclusões e trabalhos futuros

Os resultados preliminares mostraram que utilizar Redes Neurais para treinar um conjunto de dados previamente conhecidos e gerar modelos para classificar dados de queimadas de forma semiautomática é uma abordagem promissora. Ela é capaz de distinguir um grande número de falsas detecções gerados pela classificação de queimadas proposta em [Melchiori et al, 2014]. O resultado prático é uma diminuição do esforço na validação dos dados por especialistas antes da publicação oficial.

Como parte de um trabalho em andamento, estamos ampliando a área de estudo e continuamos as pesquisas em busca de novas configurações e novos atributos que

possam enriquecer o modelo. Para trabalhos futuros sugerimos a utilização de outros modelos de classificação e a incorporação de outros produtos de dados relacionados à queimadas como o risco de incêndio e focos de calor para gerar modelos que possam distinguir com maior precisão cicatrizes de queimadas e cicatrizes de falsas detecções.

References

- Boschetti, L. et al. MODIS-Landsat fusion for large area 30m burned area mapping. *Remote Sensing of Environment*, New York, v. 161, p. 27-42, Mar. 2015.
- Bowman, D.M.; Balch, J.K.; Artaxo, P.; Bond, W.J.; Carlson, J.M.; et al. Fire in the earth system. *Science*, v. 324, p. 481-484, 2009.
- Instituto Nacional de Pesquisas Espaciais (INPE). Programa de Monitoramento de Queimadas. INPE, São José dos Campos, 2017. Disponível em: <http://www.inpe.br/queimadas>. Acesso em: 01 ago. 2017.
- Li, J., & Roy, D. (2017). A Global Analysis of Sentinel-2A, Sentinel-2B and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. *Remote Sensing*, 9(9), 902.
- Libonati, R.; DaCamara, C.C.; Setzer, A.W.; Morelli, F.; Melchiori, A.E. An Algorithm for Burned Area Detection in the Brazilian Cerrado Using 4 m MODIS Imagery. *Remote Sensing*, v. 7, p. 15782-15803, 2015.
- Melchiori, A. E.; Setzer, A.W.; Morelli, F.; Libonati, R.; Cândico, P.A; Jesús, S.C. A Landsat- TM/OLI algorithm for burned areas in the Brazilian Cerrado: preliminary results. In: *Advances in Forest Fire Research, VII International Conference on Forest Fire Research*, Universidade de Coimbra, Portugal, p. 1302- 1311, 17-21/Nov//2014.
- Ministério da Ciência, Tecnologia e Inovação (MCTI). Estimativas Anuais de Emissões de Gases de Efeito Estufa no Brasil (3a edição). MCTI, Brasília, 2017. Disponível em: <http://sirene.mcti.gov.br/documents/1686653/1706227/f2a4d2ed-7478-4877-b748-1f5e32f31f60?t=1480510629875.pdf> Acesso em: 01 ago. 2017.
- Ministério do Planejamento, Orçamento e Gestão (MPOG). Plano Plurianual 2016-2019: Desenvolvimento, produtividade e inclusão social. MPOG, Brasília, 2015. Disponível em: www.planejamento.gov.br/assuntos/planeja/plano-plurianual/relatorio-objetivos.pdf. Acesso em: 12 set. 2017.
- Pereira, A. A. et al. Avaliação de índices espectrais para identificação de áreas queimadas no cerrado utilizando dados LandSat TM. *Revista Brasileira de Cartografia*, Rio de Janeiro, v. 8, n. 68, p. 1665-1680, 2016.
- Katagis, T.; Gitas, I.Z.; Toukiloglou, P.; Veraverbeke, S.; Goossens, R. Trend analysis of medium- and coarse-resolution time series image data for burned area mapping in a mediterranean ecosystem. *Int. J. Wildland Fire* 2014.
- USGS, Serviço Geológico dos Estados Unidos, Disponível em: <https://earthexplorer.usgs.gov/>. Acesso em: 01 ago. 2016.
- Venables, W. N. & Ripley, B. D. (2002) *Modern Applied Statistics with S*. Fourth Edition. Springer, New York. ISBN 0-387-95457-0.