

Identificação de pivôs centrais usando composições de bandas e um método rápido de *Deep Learning*

Denis M. de A. Eiras¹, Mikhaela A. J. S. Pletsch¹, Marcos L. Rodrigues¹, Karine R. Ferreira¹, Thales Sehn Körting¹

¹Instituto Nacional de Pesquisas Espaciais – INPE

Av. dos Astronautas, 1758 – 12227-010 – São José dos Campos – SP – Brasil

{denis.eiras, mikhaela.pletsch, marcos.rodrigues, karine.ferreira, thales.korting}@inpe.br

Abstract. *This paper presents a technique to identify central pivot (CP) using patches of images containing only one CP, composed by varied bands of the Landsat 8 OLI sensor and spectral indices, through a fast Convolutional Neural Network (CNN). Different combinations of bands and indexes were tested in this work, as infrared band and NDVI. The obtained results indicate best accuracy (95,85%) when using bands not commonly used in CNNs, surpassing some works. CNN also demonstrated advantages in terms of speed, by classifying a patch of image containing CP in only 0.28 milliseconds, revealing great potential for CP identification in remote sensing images, available in official catalogs.*

Resumo. *Este trabalho apresenta uma técnica para identificar pivô central (PC) usando partes de imagens que contém um PC, compostas por bandas variadas do sensor OLI do Landsat 8 e índices espectrais em uma Rede Neural Convolutacional (CNN) rápida. Diferentes combinações de bandas e índices foram testadas, como banda infravermelha e NDVI. Os resultados obtidos indicaram a melhor acurácia (95,85%) utilizando bandas não comumente utilizadas em CNNs, superando alguns trabalhos. A CNN também demonstra vantagens em termos de velocidade ao classificar uma imagem em 0,28 milissegundos, revelando grande potencial para identificação de PC em imagens de sensoriamento remoto, disponíveis em catálogos oficiais.*

1. Introdução

Pivô central (PC) é o sistema de irrigação mecanizada que mais cresce no país, uma vez que garante as necessidades hídricas de diferentes culturas. No Brasil, PCs são responsáveis por irrigar cerca de 20% da área irrigada total. A identificação de PCs no Brasil é de grande importância, tanto para gerir o balanço hídrico, bem como para estudar os impactos ambientais decorrentes desse tipo de prática [Fontelle, 2017]. Apesar disso, ainda não existem metodologias automatizadas para identificação de PCs em imagens de sensoriamento remoto.

Rodrigues *et al.* (2020) estudaram a identificação de PCs no Cerrado utilizando técnicas de Canny, transformada circular de Hough e séries temporais sobre índices de vegetação extraídos de produtos do MODIS, com acurácias de até 90%. Ferreira *et al.* (2011) utilizaram um método de segmentação que buscou contabilizar a área irrigada por pivôs, resultando em um índice kappa de 0,94.

Visando resultados ainda mais acurados, alguns trabalhos utilizaram as Redes

Neurais Convolucionais (ou *Convolutional Neural Networks – CNN*), utilizadas no reconhecimento de objetos em imagens, devido à sua adaptabilidade à formas e tons. A MobileNet é uma das arquiteturas de CNN que apresenta resultados similares ou superiores a outras arquiteturas mais pesadas em um tempo consideravelmente menor [Howard *et al.*, 2017]. O trabalho de Zhang *et al.* (2018) foi um dos pioneiros no uso de CNN para identificação de PCs, utilizando imagens *Red-Green-Blue* (RGB) do satélite Landsat 5. Cada PC está contido em 25 imagens, com uma pequena diferença de posição, para a identificação do centro de cada PC, resultando em uma precisão de 95,85% e um *recall* de 93,33% em uma CNN Le-Net adaptada. Albuquerque *et al.* (2020) usaram imagens do Landsat 8 para avaliar diferentes ambientes no Brasil central e mudanças sazonais. Foram testados três modelos de CNN e técnicas de reconstrução de imagem através de intervalos de sobreposição entre quadros, totalizando 10.997.161 PCs. Os melhores resultados apontaram a CNN U-Net, com F1-score e Kappa de 0,9638 e acurácia de 98,80%. Saraiva *et al.* (2020) também utilizaram a CNN U-Net, buscando velocidade de treinamento. Foram utilizadas 4 bandas do satélite PlanetScope, totalizando 42.000 imagens com resolução espacial aproximada de 3 m, levando à acurácia de 99%. Seu método levou aproximadamente 22 h para o treinamento e 10 minutos para testar cada grade de sua área de estudo, utilizando uma máquina com processador de 4 núcleos, 50 GB de RAM e uma GPU NVIDIA Tesla K80.

No seu uso mais comum, imagens de fotos são utilizadas para a classificação em CNNs. Tratando-se da identificação de PCs com CNNs na literatura atual, encontram-se trabalhos que utilizaram somente as bandas do espectro visível (RGB). Outras bandas não tradicionais, como o infravermelho, e índices espectrais, como o *Normalized Difference Vegetation Index* (NDVI) [Vermote *et al.*, 2016], que são muito utilizados no mapeamento de alvos de vegetação, poderiam substituir bandas do espectro visível para compor as imagens utilizadas pela CNN. Sendo assim, esse trabalho tem como objetivo aplicar uma técnica baseada em CNN para testar a acurácia da identificação de PCs em recortes de imagens, compostas por bandas do Landsat 8 e índices espectrais. Os resultados demonstram quais foram as melhores e piores acurácias obtidas com cada composição de bandas e os tempos de treinamento e validação.

2. Materiais e métodos

2.1. Área de estudo

A área de estudo foi selecionada com base na concentração de PCs, principalmente no Cerrado Brasileiro, exibida à esquerda e ampliada à direita da Figura 1.

2.2. Materiais

A Agência Nacional de Águas (ANA) realizou o mapeamento dos PCs ativos e inativos entre 1985 e 2017 de todo o território brasileiro [Fontenelle *et al.*, 2019]. Foram mapeados 23.181 polígonos de PCs (em azul na Figura 1) em um arquivo do tipo *shapefile*, utilizados para extrair partes de imagens de PCs. Foram utilizadas as bandas 1 a 8 do Landsat 8, cujas cenas estão localizadas com contorno em preto (Figura 1):

- A) Órbita 220 e ponto 71, de 19/09/2017, 1290 PCs: treinamento da CNN;
- B) Órbita 220 e ponto 72, de 19/09/2017, 1290 PCs: treinamento da CNN;
- C) Órbita 221, ponto 71, de 10/09/2017, 1461 PCs: validação das composições;
- D) Órbita 221 e ponto 76, de 24/07/2017, 1441 PCs: validação adicional, em região com maior diversidade de formas de polígonos;

E) Órbita 220 e ponto 69, 731 PCs: validação adicional, em outros períodos (03/11/2016; 06/01/2017; 14/05/2017; 18/08/2017).

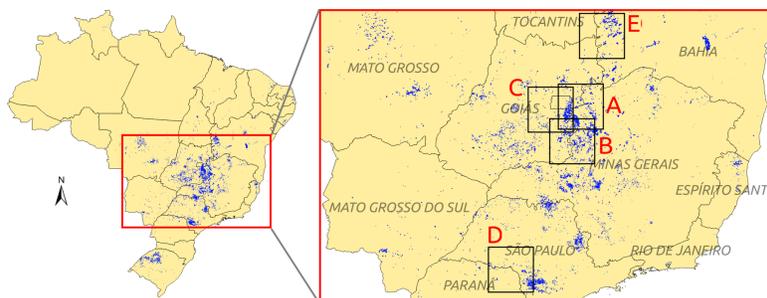


Figura 1. PCs (polígonos em azul), e localização das cenas utilizadas, no Cerrado, para treinamento (A e B) e validação (C, D e E) da CNN.

As cenas (Figura 1) foram adquiridas no catálogo de imagens do INPE (2020), selecionadas com base na grande quantidade de PCs em um mesmo bioma do Brasil (Cerrado) e baixa cobertura de nuvens, em sua maioria, com exceção de uma imagem de 06/01/2017, localizada no oeste da Bahia (E), com 8% de cobertura.

A MobileNet foi a arquitetura de CNN utilizada neste trabalho, que possui parâmetros configuráveis, como o Multiplicador de Largura, que reduz uniformemente a largura em cada camada (0% a 100%), e o parâmetro Multiplicador de Resolução (128, 160, 192 e 224), que reduz a representação interna de cada camada pelo mesmo multiplicador [Howard *et al.*, 2017]. Embora a arquitetura básica MobileNet seja pequena e de baixa latência, pode ser necessário diminuir seu tamanho e aumentar a velocidade do modelo, através da configuração destes parâmetros.

Os métodos propostos a seguir foram implementados em uma aplicação, construída com a linguagem Python v.3.7, a biblioteca TensorFlow 1.13.11 e outras bibliotecas, utilizando em um *laptop* com processador Intel i7 de sétima geração e 16 GB de memória RAM, sem a utilização de GPU.

2.3. Métodos

As seguintes etapas automatizadas da aplicação foram executadas (Figura 2):

1. geração de imagens como composição de bandas, para treinamento e validação;
2. extração de imagens com PCs e não pivôs centrais (NPC), através do mapeamento dos polígonos;
3. treinamento e testes da CNN, utilizando diferentes configurações MobileNet;
4. validação da melhor configuração da CNN, nas cenas de treinamento;
5. validação da melhor configuração da CNN, nas cenas de validação.

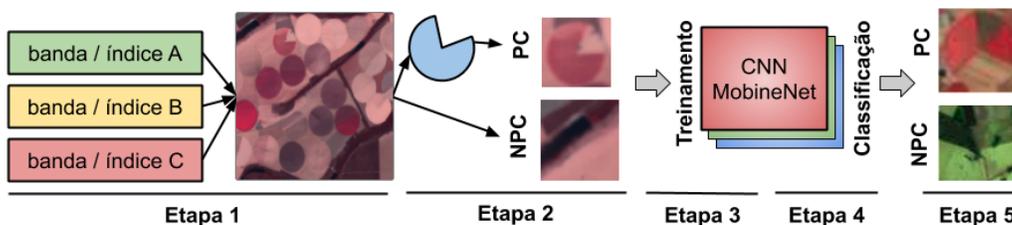


Figura 2. Etapas do método proposto.

A etapa 1 gerou uma imagem para treinamento e outra para validação, como

composição de 3 bandas ou índices espectrais, com valores normalizados e reescalados entre 0 e 255, para a redução de dimensionalidade. A etapa 2 efetuou recortes nas imagens, envolvendo os polígonos de PCs contidos no arquivo *shapefile*. O tamanho da imagem foi acrescido de uma margem de 10% do tamanho do PC, para que estes sempre estejam enquadrados na imagem. Um processo semelhante extraiu imagens contendo NPC de tamanhos aleatórios, dentro da faixa de tamanho de PC, sendo a mesma quantidade de imagens de PCs. A etapa 3 realizou diversos treinamentos como combinações dos parâmetros: arquivos com as bandas do Landsat 8, multiplicador de resolução da rede MobileNet e parâmetros para pré-processar as imagens, tais como espelhamento e brilho aleatório, a fim de aumentar a acurácia da validação. Na etapa 4, para cada treinamento gerado na etapa 3, a classificação foi realizada em 20% das imagens de treinamento. A melhor configuração de treinamento identificada com base na classificação foi armazenada para a etapa 5, a qual realizou testes com imagens de validação, excluindo-se daí os recortes de PCs utilizados no treinamento. Essa última etapa permite computar às acurácias e tempo de execução para cada composição de banda testada.

3. Resultados

O melhor resultado foi alcançado utilizando uma CNN com as seguintes configurações: Multiplicador de resolução 224, Multiplicador de Largura 1, sem uso de pré-processamento nas imagens. Essa rede foi treinada com 5160 imagens das cenas em A e em B (Figura 1), em um tempo médio de 7 minutos e 36 segundos, suficientes para atingir a máxima acurácia na validação. A validação foi executada em 2922 recortes da cena C (Figura 1) em aproximados 0,42 segundos (0,28 milissegundos por recorte, em média), onde se considerou somente o tempo do teste de cada imagem na memória.

Dado o conjunto de 35 combinações possíveis das bandas utilizadas, temos na Tabela 1 sumarizados nas 5 primeiras linhas os melhores resultados e nas 5 últimas os piores, em termos de acurácia estimada através da validação. A precisão e o *recall* foram calculados sobre a classe PC. A composição RGB (4, 3, 2) ficou apenas na 27ª posição entre as melhores acurácias (92,84%). Para comparação, a Tabela 2 apresenta os melhores resultados obtidos com a combinação de índices espectrais e bandas do sensor OLI. De maneira a demonstrar a generalização alcançada, foram utilizadas imagens de outras regiões e períodos, usando as bandas 4, 6 e 7. A Tabela 3 apresenta os resultados obtidos nas posições D e E (Figura 1).

Tabela 1. Melhores e piores acurácias na validação das composições

Composição	Acurácia (%)	Precisão (%)	Recall (%)	F1-Score	Kappa
Bandas 4, 6, 7	95,85	96,91	94,72	0,9580	0,9585
Bandas 3, 5, 6	95,45	97,36	93,42	0,9535	0,9544
Bandas 5, 6, 7	94,96	94,81	95,13	0,9497	0,9496
Bandas 2, 4, 5	94,90	94,50	95,35	0,9492	0,9490
Bandas 2, 6, 7	94,76	96,05	93,36	0,9469	0,9476
Bandas 4, 5, 6	91,99	98,12	85,62	0,9144	0,9198
Bandas 1, 5, 6	91,51	98,48	84,32	0,9085	0,9150
Bandas 1, 3, 4	90,97	98,54	83,16	0,9020	0,9096
Bandas 1, 2, 3	90,52	95,68	84,87	0,8995	0,9051
Bandas 1, 2, 5	81,86	99,47	64,07	0,7794	0,8185

Tabela 2. Validação de composições com índices espectrais e bandas

Índices+bandas	Acurácia (%)	Precisão (%)	Recall (%)	F1-Score	Kappa
NDVI, 6, 3	94,10	92,93	95,47	0,9418	0,9410
SAVI, 6, 7	93,15	92,00	94,52	0,9324	0,9315
NDVI	92,02	89,05	95,82	0,9231	0,9202
SAVI	91,20	87,95	95,48	0,9156	0,9119
EVI	88,26	84,80	92,32	0,8816	0,8825

Tabela 3. Validações em outras cenas e períodos

Cena, data	Acurácia (%)	Precisão (%)	Recall (%)	F1-Score	Kappa
E, 18/08/2017	95,69	95,13	96,30	0,9571	0,9568
E, 14/05/2017	94,66	99,24	90,01	0,9404	0,9465
E, 06/01/2017	89,67	100,00	79,34	0,8848	0,8965
E, 03/11/2016	88,61	99,47	77,67	0,8723	0,8861
D, 24/07/2017	75,53	97,79	52,25	0,6811	0,7552

4. Discussão

A CNN deste e outros trabalhos superaram resultados obtidos por técnicas baseadas em detectores de bordas, círculos e segmentação, utilizados nos trabalhos de Rodrigues *et al.* (2020) e Ferreira *et al.* (2011), devido à sua adaptabilidade à diversidade de formas. No entanto, as CNN requerem uma grande quantidade de amostras para treinamento, que devem ser extraídas de regiões e períodos próximos às amostras de validação.

A utilização de índices não apresentou melhores resultados que a utilização exclusiva de bandas, pois a CNN encontrou parâmetros ótimos que superaram índices. De outro lado, a utilização de diferentes bandas melhorou a acurácia da CNN com relação às bandas RGB. As bandas 6 e 7 (infravermelho de ondas curtas) e 5 (infravermelho próximo) apareceram mais vezes nas 10 melhores composições (Figura 3), superando acurácias do trabalho de Zhang *et al.* (2018), que utilizou bandas RGB.

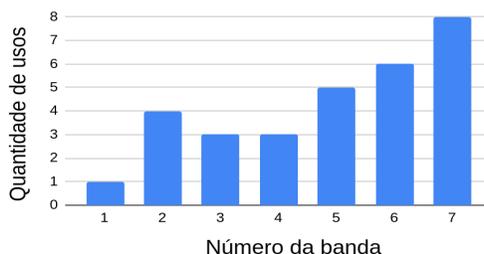


Figura 3. Quantidade de usos das bandas nas 10 melhores composições.

Albuquerque *et al.* (2020) usaram uma técnica de reutilização de exemplos, e Saraiva *et al.* (2020) imagens de melhor resolução, superando a acurácia desse trabalho em, respectivamente, 3,03% e 3,15%, mas utilizando CNN de arquiteturas maiores e dedicadas. Saraiva *et al.* executaram o treinamento em 22 h, muito superior ao tempo da técnica proposta neste trabalho, que levou 7 minutos para o treinamento, demonstrando um ótimo balanceamento entre qualidade e velocidade de identificação de PCs.

5. Conclusões

As melhores composições foram atingidas ao utilizar as bandas 6 e 7, sendo que a melhor composição (95,85%) utilizou as bandas 4, 6 e 7. A validação de uma cena mais distante (Figura 1-D) piorou a acurácia em aproximados 20%, indicando que há características distintas de vegetação, forma e tamanho de PC. A validação de cenas próximas em diferentes períodos (Figura 1-E) apresentou uma variação de acurácia de até 7,8%. Estas validações sugerem que é preciso realizar treinamentos em localidades e períodos mais próximos da cena a ser validada para se atingir melhores resultados.

Em trabalhos futuros será desenvolvida um CNN direcionada para identificar e quantificar PCs, aproveitando o potencial demonstrado pelo uso das composições de bandas não RGB, além de utilizar mais de 3 bandas como entrada, possibilitando a rápida identificação de PCs em catálogos oficiais de sensoriamento remoto.

Agradecimentos

Os autores agradecem aos Prof. Dr. Gilberto R. de Queiroz, Dra. Lúbia Vinhas e Dra. Karine R. Ferreira e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), processos 140377/2018-2 e 303360/2019-4. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Referências

- Albuquerque, A. O. de, de Carvalho Júnior, O. A., Carvalho, O. L. F. D. & Fontes Guimarães, R. (2020). Deep semantic segmentation of center pivot irrigation systems from remotely sensed data. *Remote Sensing*, 12(13), 2159.
- Ferreira, E., Dantas, A. A. A., de Toledo, J. H. (2011). Classificação de áreas irrigadas por pivôs centrais utilizando como base a segmentação. *Irriga*, 16(2), 145-152.
- Fontelle, T., Ferreira, D., Guimarães, D., & Landau, E. (2019). Levantamento da agricultura irrigada por pivôs centrais no Brasil. Livro científico (ALICE).
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications.
- INPE – Instituto Nacional de Pesquisas Espaciais (2020). www.dgi.inpe.br/catalogo
- Rodrigues, M. L., Körting, T. S., de Queiroz, G. R., & da Silva, L. A. R. (2020). Detecting Center Pivots In Matopiba Using Hough Transform And Web Time Series Service. In 2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS), 189-194.
- Saraiva, M., Protas, É., Salgado, M., & Souza Jr, C. (2020). Automatic Mapping of Center Pivot Irrigation Systems from Satellite Images Using Deep Learning. *Remote Sensing*, 12(3), 558.
- Vermote, E., Justice, C., Claverie, M., & Franch, B. (2016). Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product. *Remote Sensing of Environment*, 185, 46-56.
- Zhang, C., Yue, P., Di, L., & Wu, Z. (2018). Automatic identification of center pivot irrigation systems from landsat images using convolutional neural networks. *Agriculture*, 8(10), 147.