

Mapeamento Sistemático sobre Redes Neurais Convolucionais Aplicadas na Segmentação de Imagens Aéreas

Fabricio Bizotto¹, Mauren Louise Sguario Coelho De Andrade¹

*^aDepartamento Acadêmico de Informática
Programa de Pós-Graduação em Ciência da Computação (PPGCC)
Universidade Tecnológica Federal do Paraná (UTFPR)
Ponta Grossa, Brasil
^bfabriciobizotto@alunos.utfpr.edu.br, mlsguario@utfpr.edu.br*

Resumo

Existem muitos estudos atuais publicados sobre Redes Neurais Convolucionais (CNN) em diferentes contextos. Foi feito um mapeamento sistemático sobre a utilização de redes CNN aplicadas na segmentação de imagens aéreas. A partir do levantamento inicialmente feito, 5 artigos foram selecionados e analisados. Como resultado, percebe-se que essa linha de pesquisa vem crescendo, mas seria necessário analisar mais publicações para responder melhor às questões de pesquisa. Entre as redes CNN utilizadas, cada uma conseguiu contribuir de alguma forma na segmentação de imagens aéreas.

Abstract

There are many current studies published on Convolutional Neural Networks (CNN) in different contexts. A systematic mapping of the use of CNN networks applied in the segmentation of aerial images was carried out. From the survey initially carried out, 5 articles were selected and analyzed. As a result, it is clear that this line of research has been growing, but it would be necessary to analyze more publications to better answer the research questions. Among the CNN networks used, each managed to contribute in some way to the segmentation of aerial images.

Palavras-chaves: Rede Neural Convolucional, Imagens Aéreas, Segmentação Semântica

1. Introdução

O desenvolvimento de novas abordagens em Machine Learning (ML) proporciona aos pesquisadores alternativas para a resolução de problemas gerais, através de modelos computacionais criados e treinados a partir de uma grande quantidade de dados. Uma das abordagens que está ganhando espaço é a Rede Neural Convolucional (CNN), oriunda das áreas de processamento de imagens e visão computacional. O algoritmo possibilita vantagens em comparação aos métodos anteriores de última geração, superando nas tarefas de

classificação de imagens, reconhecimento e localização de objetos e rotulagem semântica de pixels[1].

A primeira implementação da CNN foi feita por LeCun [2]. O experimento foi aplicado no reconhecimento de dígitos manuscritos e resultou em uma precisão de 99,2% de acerto na identificação dos números. O processamento de imagens utilizando CNN cresceu nos anos seguintes. São utilizadas em diversos contextos. No contexto relacionado a imagens aéreas, as redes CNN são utilizadas, por exemplo, para o mapeamento aéreo de piscinas [3], na avaliação dos danos sísmicos em edifícios, para melhorar a resposta a emergências em locais sujeitos a terremotos [4], para detectar objetos de interesse em imagens aéreas capturadas por UAVs [5] e também para identificar e classificar ervas daninhas que crescem no meio do soja [6]. Grandes repositórios de imagens tais como ImageNet surgiram e a otimização no desempenho das unidades de processamento gráfico (GPU) contribuíram para o sucesso das CNN. Além disso, competições como ImageNet Challenge [7] e Visual Object Classes Challenge [8], publicadas na revista International Journal of Computer Vision, atraem pesquisadores. Resultados desse tipo de competição gerou modelos de rede CNN, tais como AlexNet [9] e VGG-Net [10].

Com isso, a segmentação de imagens, cujo objetivo é separar a área/objeto de interesse de seu fundo, sofreu grande avanço com a aplicação de segmentação semântica. A segmentação semântica é caracterizada por classificar objetos que possuem algum rótulo, conhecido ou desconhecido. [11] Essa técnica pode ser utilizada em diferentes contextos, desde o sensoriamento remoto por satélite [12], análise de vídeos de embriões em desenvolvimento [13] e interação humano-computador [11].

Dentre as áreas de aplicação de segmentação semântica citadas, a segmentação e rotulagem de imagens aéreas tem possibilitado o acompanhamento e prevenção de problemas relacionados ao desmatamento e apropriação de áreas de preservação em florestas em diferentes regiões do mundo [14]. Neste sentido, o desenvolvimento de uma nova metodologia de segmentação semântica com uso de CNN capaz de aprimorar o reconhecimento e rotulagens de áreas de preservação poderá contribuir com o trabalho das autoridades ambientais na fiscalização e demarcação de áreas já afetadas.

Portanto, com o objetivo de responder algumas questões relacionadas à aplicação de CNN em imagens aéreas, este trabalho apresenta um mapeamento sistemático, o qual irá identificar e mapear quais abordagens têm sido utilizadas na segmentação semântica de imagens aéreas com o uso de CNN. A ideia principal é coletar informações a fim de identificar estratégias para o aprimoramento dos resultados de classificação e segmentação de áreas de preservação brasileiras, em especial na área de preservação da Mata Atlântica na região de Petrópolis, Rio de Janeiro. Este artigo está estruturado da seguinte maneira: Após esta breve introdução, na seção 2 é descrito o método de pesquisa. Na seção 3, os resultados são discutivos e na seção 4 uma breve conclusão.

2. Método de Pesquisa

Para levantar o estado da arte sobre a utilização de redes CNN no contexto de imagens aéreas, foi feito um mapeamento sistemático seguindo os procedimentos essenciais propostos por Petersen [15], as quais serão descritos a seguir:

2.1. Metodologia

O mapeamento foi conduzido entre abril e maio de 2021 e cobriu publicações sobre utilização de redes CNN em imagens aéreas, no período de 2017 a 2021.

2.2. Questões de Pesquisa

Foram definidas questões que considerem pontos relevantes a serem estudados:

- QP1: Quais e como as redes CNN são utilizadas para classificação de imagens aéreas?
- QP2: Em quais contextos as redes CNN foram aplicadas?
- QP3: Quais redes CNN funcionam melhor para classificação de imagens aéreas?

2.3. Palavras chave e string de busca

As palavras chave estão em inglês e foram definidas de acordo com o tema pesquisado. São elas: CNN, Convolutional Neural Network, Image Segmentation, Image Classification, Aerial Image e Aerial Imagery. A string de busca foi criada a partir da combinação lógica entre elas, e ficou da seguinte forma: (“CNN” OR “Convolutional neural network”) AND (!Image Segmentation” OR “Image Classification”) AND (“Aerial image” OR “Aerial Imagery”).

2.4. Critérios de inclusão e exclusão

Para refinar a pesquisa, os seguintes critérios de inclusão de artigos foram aplicados: (i) artigos escritos em inglês; (ii) artigos com relevância ao assunto abordado; (iii) revistas com relevância na área; (iv) artigos com qualis maior ou igual a B1. Também foram aplicados os seguintes critérios de exclusão: (i) artigos duplicados; (ii) artigos com título e resumo fora do assunto abordado.

2.5. Bases de Pesquisa

As máquinas de busca utilizadas foram: *ACM Digital Library*, *IEEE Xplore*, *Scopus* e *Science Direct*. A string de busca utilizada foi a mesma em todas as bases. A pesquisa resultou em 837 artigos não duplicados. A maior parte dos estudos é recente e sugere um crescente interesse pelo tema nos últimos anos, devido a fatores como avanço em termos de hardware e economia de tempo, aliados a resultados promissores em termos de segmentação. Foram selecionados 5 artigos para serem analisados.

3. Resultados e Discussão

Respondendo a QP1, cada aplicação faz o pré-processamento da rede, normalizando os dados de entrada e separando as imagens em grupos para a fase de treinamento, validação e teste. Cada estrutura de rede CNN é configurada com uma quantidade diferente de camadas de convolução. A Figura 1 mostra os *datasets* e as redes CNN utilizadas em cada trabalho.

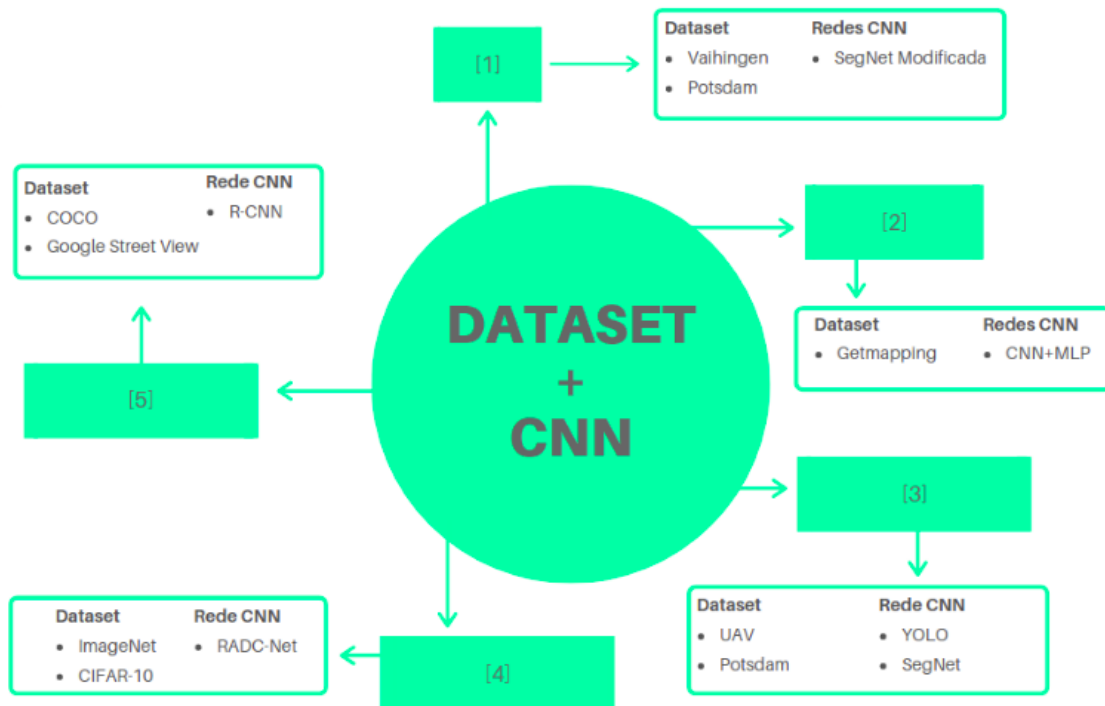


Figura 1: *Datasets* e redes CNN utilizadas

Cada estudo utilizou uma rede CNN diferente e *datasets* com imagens aéreas. O *dataset* mais utilizado foi da cidade de Potsdam, Alemanha. Respondendo a QP2, diferentes contextos foram abordados nos trabalhos: prevenção de risco de incêndio; área agrícola; mapeamento de árvores. As questões anteriores colaboram para responder a QP3. Contudo, não é possível definir qual rede CNN funciona melhor para classificar imagens aéreas. Cada implementação auxiliou de alguma forma, seja para reduzir o erro de validação, para reduzir problemas de overfitting ou para identificação de árvores com alta taxa de precisão.

4. Conclusão

O foco deste estudo de mapeamento sistemático é investigar como as redes CNN estão sendo aplicadas para segmentação semântica no contexto de imagens aéreas. Apesar de terem sido mapeados 837 artigos, apenas 5 foram utilizados. Mais artigos precisam ser

considerados para enriquecer o mapeamento e obter mais dados para responder melhor às questões de pesquisa elaboradas. O que se conclui desse mapeamento sistemático é que a grande quantidade de publicações recentes em jornais e revistas demonstra o crescente interesse pelo tema.

Referências

- [1] K. Nogueira, O. A. B. Penatti, A. Santos, Towards better exploiting convolutional neural networks for remote sensing scene classification, *Pattern Recogn.* 61 (2017) 539–556.
- [2] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, *Proc. IEEE* 86 (11) (1998) 2278–2324.
- [3] B. Lima, L. Ferreira, J. Moura, Helping to detect legal swimming pools with deep learning and data visualization, *Procedia Computer Science* 181 (2021) 1058–1065.
- [4] C. Xiong, Q. Li, X. Lu., Automated regional seismic damage assessment of buildings using an unmanned aerial vehicle and a convolutional neural network, *Automation in Construction* 109 (2020) 102994.
- [5] L. Zhaokun, L. X. Zhao, Y. Liu, B. Huang, Z. H. Richang, A lightweight multi-scale aggregated model for detecting aerial images captured by uavs, *Journal of Visual Communication and Image Representation* 77 (2021) 103058.
- [6] A. S. Ferreira, D. M. Freitas, G. G. Silva, H. Pistori, M. T. Folhesc., Weed detection in soybean crops using convnets, *Computers and Electronics in Agriculture* 143 (2017) 314–324.
- [7] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, C. Berg, L. Fei-Fei, Imagenet large scale visual recognition challenge, *Int. J. Comput. Vision* 115 (3) (2015) 211–252.
- [8] M. Everingham, S. M. A. Eslami, L. V. Gool, C. K. I. Williams, J. Winn, A. Zisserman, The pascal visual object classes challenge: A retrospective, *Int. J. Comput. Vision* 111 (1) (2015) 98–136.
- [9] A. Krizhevsky, I. Sutskever, G. E. Hinton., Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems* 25 (2012) 1097–1105.
- [10] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, *Computer Vision and Pattern Recognition* (2015).

- [11] G. Yanming, L. Yu, T. Georgiou, M. S. Lew, A review of semantic segmentation using deep neural networks, *International Journal of Multimedia Information Retrieval* 7 (2018) 87–93.
- [12] M. Coffe, B. Schaeffer, R. Zimmerman, V. Hill, J. Li, K. Islam, P. Whitman, Performance across worldview-2 and rapideye for reproducible seagrass mapping, *Remote Sensing of Environment* 250 (2020) 112036.
- [13] F. Ning, D. Delhomme, Y. LeCun, F. Piano, L. Bottou, P. Barbano, Toward automatic phenotyping of developing embryos from videos, *IEEE Transactions on Image Processing* 14 (9) (2005) 1360–1371.
- [14] A. A. Sant’Anna, L. Costa, Environmental regulation and bail outs under weak state capacity: Deforestation in the brazilian amazon, *Ecological Economics* 186 (2021) 107071.
- [15] K. Petersen, S. Vakkalanka, L. Kuzniarz, Guidelines for conducting systematic mapping studies in software engineering, *Information and Software Technology* 64 (2015) 1–18.