**Análise do Artigo “Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery”**1

Sílvia Mourão

(fc57541@alunos.fc.ul.pt)

**RESUMO**

Este trabalho foca-se na obtenção de mapas de ocupação do solo em zonas urbanas. Neste estudo é proposta uma abordagem com um método de “semi-transfer deep convolutional neural network” (STDCNN) como resposta a alguns problemas dos algoritmos existentes atualmente. Além disso é também considerada uma abordagem de decomposição em esqueleto para manter a integridade dos padrões de ocupação do solo. Foram realizados dois casos de estudo, utilizando uma imagem WorldView3 de Hong Kong e uma imagem WorldView2 de Shenzhen. Os resultados demonstraram uma Overall Accuracy (OA) de 91.25% para a região de Hong Kong e 80% para a região de Shenzhen, sendo que foram também registadas melhorias nos mapas de ocupação do solo resultantes pela decomposição em esqueleto quando comparada com a decomposição uniforme.

**1. Introdução**

O mapeamento de ocupação do solo urbano é um método fundamental para reconhecer e localizar utilizações do solo para diferentes propósitos, como monitorização de ambientes urbanos, planeamento e ordenamento do território. No passado recente, os métodos utilizados para atualizar estes mapas dependiam da interpretação de fotografia aérea e trabalho de levantamento em campo, ambos dispendiosos e morosos. Com o desenvolvimento de tecnologias de deteção remota, um grande número de imagens de elevada resolução espacial que cobrem uma área urbana podem ser obtidas. Para classificar estas imagens, utilizaram-se métodos de classificação baseados no pixel, em objetos 2, em grafos de vizinhança 3, entre outros. No entanto, estes métodos têm limitações, pois as imagens de alta resolução espacial são demasiado complexas para serem caracterizadas ao nível do pixel, estão muito dependentes do sistema de classificação escolhido ou requerem conhecimento prévio sobre a cobertura do solo.

“Deep learning” é um processo através do qual um conjunto de algoritmos de aprendizagem automática tentam modelar abstrações de dados de alto nível através da utilização de uma arquitetura profunda composta por múltiplas transformações não lineares. Entre as várias técnicas de “deep learning”, o método “deep convolutional neural network”, ou DCNN, tem conseguido atingir um elevado nível de desempenho na classificação de ocupação do solo. Os DCNN são compostos por múltiplas camadas de convolução e são capazes de aprender características abstraídas dos valores de pixel originais das imagens de ocupação do solo. No entanto, o aumento no número de camadas aumenta o número de parâmetros e requer um elevado número de amostras de treino. Para reduzir esse número, alguns investigadores propuseram a utilização de “transfer DCNNs” ou de DCNNs pequenos. No entanto, os “transfer DCNNs” não se adaptam a imagens multiespectrais e os DCNN pequenos são incapazes de aproveitar as vantagens da arquitetura profunda. Além disso, os métodos existentes de classificação tendem a dividir imagens de tamanho grande em pequenas unidades de processamento de tamanhos fixos através do método de decomposição uniforme, o que pode prejudicar a classificação pois pode dividir padrões de ocupação do solo.

Para resolver este problema, este estudo propôs um método de “semi transfer convolutional neural network” (STDCNN) para a classificação do solo urbano baseado em imagens de alta resolução espacial multiespectrais, composto por três partes: um DCNN transferido, um DCNN pequeno, e uma última parte que junta as duas. De forma a melhorar a aplicação nas imagens, foi também considerado uma a decomposição da imagem em esqueleto, que consiste num método adaptativo para divisão de unidades de mapeamento em unidades de processamento com formas regulares, o que mantem uma melhor integridade das unidades do que o método de decomposição uniforme.

**2. Área de Estudo**

Uma das áreas de estudo consiste na área metropolitana da região de Hong Kong, para a qual foram adquiridas duas imagens Worldview-3 em Outubro de 2015, com resolução espacial 1.24m e contendo 8 bandas entre 400-1040nm (coastal, azul, verde, amarelo, vermelho, rededge, NIR1 e NIR2) e ainda dados vetoriais do departamento de planeamento de Hong Kong. A outra área de estudo é a região de Shenzhen, na China, para a qual foi adquirida uma imagem Worldview-2 em Abril de 2015, com resolução espacial de 0.5m e 4 bandas espectrais (azul, verde, vermelho e IR), complementados por dados vetoriais do planeamento urbano do município de Shenzhen.

Na área de Hong Kong foram identificados 11 tipos de estruturas que podem ser utilizadas para caracterizar diferentes classes de ocupação do solo. Na área de Shenzhen, devido à maior resolução espacial da imagem worldview-2 e a diferentes condições geográficas e culturais, foram também identificadas 11 classes, sendo que algumas são partilhadas com Hong Kong e outras são exclusivas de Shenzhen. Estas classes são visíveis na tabela 1.

Tabela 1 - Classes para classificação do solo, H corresponde a Hong Kong e S corresponde a Shenzhen

|  |  |  |
| --- | --- | --- |
| Structure | Main land-use | Urban function |
| H1, S1 | Commercial | Catering, hotel, shop, residential, office |
| H2, S2 | Institutional | Education, church, scientific research, office |
| H3 | Port | Port, ship, water |
| H4, S3 | Dense residential (D\_Resid) | Residential with dense population |
| H5, S4 | Sparse residential (S\_Resid) | Village, villa, shop, old residential |
| H6, S5 | Woodland | Woods, grass, park building |
| H7 | Water | Water, ship |
| H8, S6 | Open space | Park, sports field, leisure square |
| H9, S7 | Vacant | Unused or developing area |
| H10, S8 | Industrial | Factory, warehouse |
| H11, S9 | Container terminal (C\_Termi) | Container, terminal square |
| S10 | Medium residential (M\_Resid) | Residential with medium dense population |
| S11 | Road | Highway, overpass, and intersection |

Para Hong Kong existem 36 amostras de treino para cada classe, com uma dimensão padrão de 256x256 pixels e as amostras de teste variam entre 15 e 25 para cada classe, com a mesma dimensão. Para Shenzhen existem 40 amostras de treino e 30 amostras de teste por classe, também com a mesma dimensão.

**3. Metodologia**

Para produzir um mapa de ocupação do solo a partir de uma imagem multiespectral de elevada resolução espacial foi proposto um método de mapeamento baseado em STDCN. A estrutura do STDCNN proposto inclui três partes. Em primeiro lugar um DCNN que é transferido do modelo AlexNet, já treinado por um grande set de imagens naturais (1.2 milhões de imagens e 1000 classes de características topográficas), disponível gratuitamente na internet, que permite ao STDCNN proposto adquirir uma arquitetura profunda. A segunda parte do método é um DCNN pequeno, desenhado para interpretar imagens multiespectrais. A terceira parte contém uma camada “fully connected” e uma camada “softmax”– o primeiro junta as duas partes e o segundo cria um vetor final de confiança da imagem de ocupação do solo. O STDCNN proposto inclui uma arquitetura multicamada treinável, que contem um numero de camadas convolucionais (aplicam filtros de convolução, onde o filtro tem dimensão menor que a imagem), “pooling” (extrai valores máximos ou médios de cada região espacial) e “fully connected” (semelhante ao convolucional mas utiliza filtros de dimensão igual à imagem). É ainda considerada a loss function 4, que é uma medida de quão bom é o modelo de predição em termos de ser capaz de prever o valor esperado. O STDCNN é aperfeiçoado através de otimização pelo cálculo do gradiente SGD de forma a encontrar o mínimo desta função.

Devido à grande dimensão da imagem, um método de decomposição deve ser utilizado para a sua divisão em pequenas unidades de processamento. Outros autores propuseram a divisão das imagens em parcelas baseadas em blocos de ruas. No entanto, as formas destas unidades de mapeamento são frequentemente irregulares, ou têm escalas diferentes, o que pode torná-las inaceitáveis como input no modelo DCNN. Este estudo propõe o uso de um método de decomposição baseado em esqueleto, que faz uma amostragem de cada unidade de mapeamento e representa-a com um conjunto de unidades de processamento de tamanho regular. O processo está exemplificado na figura 1.

Diagram

Description automatically generated

Figura 1 - Esquema da metodologia

As unidades de processamento geradas pelo método de decomposição podem então ser classificadas pelo STDCNN treinado. Finalmente, as unidades de processamento são combinadas para gerar um mapa de ocupação do solo.

**4. Resultados**

Com o objetivo de avaliar a classificação obtida pelo algoritmo anterior, foi gerada uma confusion matrix5 para cada uma das áreas em estudo, assim como uma confusion matrix da aplicação de outros métodos propostos por outros autores para a mesma área de estudo. Os resultados aparecem comparados na tabela 2:

Tabela 2 - Resultados da aplicação dos métodos de classificação

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Área |  | Yang and Newsam (2011) 6 | Zhao et al. (2013) 7 | Transfer DCNN (Zhao et al., 2013) 7 | Small DCNN | STDCNN |
| Hong Kong | OA (%) | 80.00 | 77.08 | 87.91 | 85.83 | 91.25 |
| Kappa | 0.779 | 0.747 | 0.867 | 0.843 | 0.903 |
| Shenzhen | OA (%) | 71.82 | 70.00 | 71.52 | 76.06 | 80.00 |
| Kappa | 0.690 | 0.670 | 0.687 | 0.737 | 0.780 |

As confusion matrixes sugerem que a aplicação do STDCNN proposto no dataset de Hong Kong atingem precisões acima de 80% para todas as classes, ultrapassando os 90% em seis das 11 classes. Para o dataset de shenzhen os resultados são ligeiramente piores, no entanto ainda assim existe uma precisão acima de 90% para 5 classes em 11, sendo que a principal dificuldade é a não obtenção de resultados satisfatórios na classificação de zonas comerciais, industriais e residenciais de média densidade devido à dificuldade em discriminar entre estas classes de ocupação do solo sem dados auxiliares.

Depois do treino, o modelo STDCNN é utilizado para classificar as unidades de processamento obtidas através da decomposição uniforme e da decomposição baseada em esqueleto. No geral, o mapa de ocupação do solo obtido utilizando o método de decomposição por esqueleto proposto são mais detalhados do que aqueles obtidos através do método de decomposição uniforme, como pode ser visível na figura seguinte. Assim, o método proposto de decomposição baseada em esqueleto fornece um melhor desempenho do que o método utilizado mais frequentemente, como pode ser visto na figura 2.

Map

Description automatically generated

Figura 2 - Resultados comparativos dos dois métodos de decomposição da imagem

**Referências Bibliográficas**

(1) Huang, B.; Zhao, B.; Song, Y. Urban Land-Use Mapping Using a Deep Convolutional Neural Network with High Spatial Resolution Multispectral Remote Sensing Imagery. *Remote Sensing of Environment* **2018**, *214*, 73–86. https://doi.org/10.1016/j.rse.2018.04.050.

(2) Blaschke, T.; Hay, G. J.; Kelly, M.; Lang, S.; Hofmann, P.; Addink, E.; Queiroz Feitosa, R.; van der Meer, F.; van der Werff, H.; van Coillie, F.; Tiede, D. Geographic Object-Based Image Analysis - Towards a New Paradigm. *ISPRS Journal of Photogrammetry and Remote Sensing* **2014**, *87*, 180–191. https://doi.org/10.1016/J.ISPRSJPRS.2013.09.014.

(3) Voltersen, M.; Berger, C.; Hese, S.; Schmullius, C. Object-Based Land Cover Mapping and Comprehensive Feature Calculation for an Automated Derivation of Urban Structure Types at Block Level. *Remote Sensing of Environment* **2014**, *C* (154), 192–201. https://doi.org/10.1016/J.RSE.2014.08.024.

(4) *Understanding Loss Functions in Machine Learning | Engineering Education (EngEd) Program | Section*. https://www.section.io/engineering-education/understanding-loss-functions-in-machine-learning/ (accessed 2022-05-14).

(5) *Classification accuracy assessment. Confusion matrix method*. http://www.50northspatial.org/classification-accuracy-assessment-confusion-matrix-method/ (accessed 2022-05-14).

(6) Yang, Y.; Newsam, S. Spatial Pyramid Co-Occurrence for Image Classification. *Proceedings of the IEEE International Conference on Computer Vision* **2011**, 1465–1472. https://doi.org/10.1109/ICCV.2011.6126403.

(7) Zhao, B.; Zhong, Y.; Zhang, L. Scene Classification via Latent Dirichlet Allocation Using a Hybrid Generative/Discriminative Strategy for High Spatial Resolution Remote Sensing Imagery. *http://dx.doi.org/10.1080/2150704X.2013.858843* **2013**, *4* (12), 1204–1213. https://doi.org/10.1080/2150704X.2013.858843.