Convolutional Neural Networks as Context-Scraping Tools in Architecture and Urban Planning

mgr inż. arch.
TOMASZ DZIEDUSZYŃSKI
Politechnika Warszawska
Wydział Architektury
ORCID: 0000-0003-2571-5128

These new, convolutional city users could roam the abstract, digitized spaces of our cities to provide insight into the architectural and urban contexts relevant to design and management processes. This article presents the results of a query of the state-of-the-art applications of Convolutional Neural Networks as architectural "city scrapers" and proposes a new, experimental framework for utilization of CNNs in context scraping in urban scale.

Data scraping

"Data scraping" or "screen scraping" is understood as an automated process of extracting useful data from a chosen document, website or a program [1]. Usually the data is obtained from a human readable output and organized into an ordered, machine-readable database [2]. Originally, in the 1970s and 1980s, screen scraping tools would capture emulated screen or terminal data provided by older mainframes [3]. Nowadays, data scraping is mostly conducted online. The first example of an automated tool, designed to autonomously roam the Internet in order to collect data on its size, was the World Wide Web Wanderer written by Matthew Gray in 1993 [4]. Since then, the Internet was no longer exclusively a human domain. Human users started sharing the World Wide Web with artificial, digital agents. Currently, 64% of all Internet traffic in the World is generated by bots [5]. Most of them use the easily accessible Application Programming Interfaces (APIs) dedicated to machine use, but some still rely on the traditionally understood data scraping.

Similarly to the Internet, our cities were originally designed with human users in mind. With the current trend of making our cities accessible online in the name of "Smart-City" or "Internet-of-Things" movements, automated urban data scraping is slowly becoming a reality. For example, within the default, Web-related meaning of the term, data scraping is used to gauge and monitor the spatial patterns of property prices extracted from real estate market websites [6]. However, the data scraping doesn't have to be bound just to the Internet-accessible data. One other way in which this offline "city scraping" can be conducted is through the use of Deep Convolutional Neural Networks (CNNs) [7-9] directly probing digitized urban spaces.

Convolutional city scraping: state of the art

Neural Networks are algorithmic tools capable of transforming a variety of input data into the desired output data by passing it through a sequence of interconnected, artificial neurons arranged in layers. In the classical, Fully-connected Neural Networks (FNNs), all the neurons within a layer are connected with all of the neurons or inputs in the previous layer. In the Convolutional Neural Networks, however, some of the fully-connected layers are replaced by convolutional filters (or kernels), which are connected only to a fraction of the previous layer's neurons and cover the whole space of the given input by sliding over it in a sequence. Thanks to this solution, one filter can detect many similar features across the whole input space [8]. CNNs were originally created for synthetic image processing tasks, such as pattern [7] and object recognition [8] and their classification [9]. Due to their capability to successfully extract features from multidimensional, spatial data, CNNs are being widely adopted in city data extraction. From the urban perspective, the most common applications can be divided into five categories: classification, semantic segmentation, monitoring, evaluation and prediction.

In the category of classification, the collected urban sensory data can be scanned and distributed across a predefined set of classes. For example, hand-taken photos of buildings can be categorized by their style and function [10] or certain urban elements like roads, buildings [11] as well as individual tree species [12] can be identified in aerial drone footage. Similarly to classification problems, the identified categories can be spatially superimposed on top of the input sensory data, segmenting it semantically. For instance, changes in urban tissue represented in aerial footage, such as new or demolished buildings, can be highlighted by a network to help in urban management [13]. Also, a convolutional network can identify high-level urban structures (roads, powerline towers, buildings, etc.) and represent them in a simplified way utilizing abstract feature maps, which could facilitate subsequent processing of the data [14]. CNNs can be useful in monitoring of urban processes, like short term changes to the urban tissue visible in satellite imagery [15] or even wear and tear of individual pavement sections [16]. The monitored processes and phenomena can then be evaluated in order to quantitatively or qualitatively assess the relevant urban contexts. Among the most interesting applications of CNNs within this category – the subjective quality, safety, desirability and attractiveness of urban space can be evaluated based on the provided pictures taken on site [17]. A research of similar approach, but limited only to the assessment of quality and activeness of street frontages, also demonstrated satisfactory results [18]. Finally, in the last class of problems, convolutional neural networks can be used to predict dynamic and cyclic phenomena within a city based on

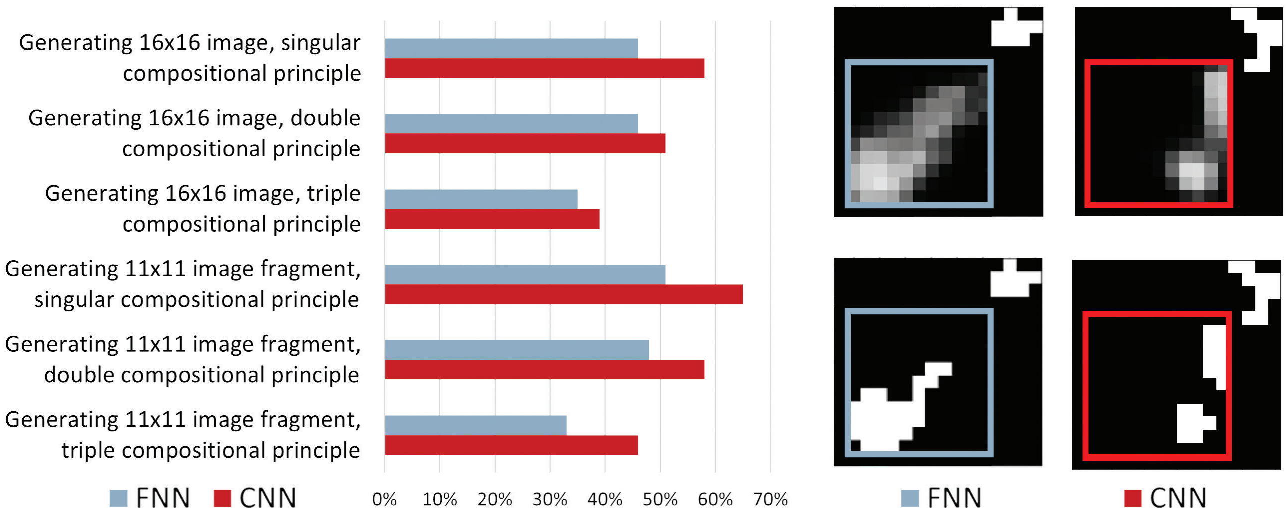

Fig. 1. Comparison of the performance (%) of FNNs and CNNs in the processing of multidimensional spatial tasks (left) and examples of the solutions generated by both types of networks (right). The networks had to learn to generate three islands fulfilling a series of compositional rules based on the given training examples [22]. A CNN learns to properly generate the islands 4-14% better than an FNN of a comparable size.

historical and present data. These applications range from traffic congestion prediction in the local [19] and citywide scale [20] to the expected urban expansion of the whole metropolises [21].

Most of these examples, however, contrary to the spirit of data scraping, rely on their own, customized sensory data, like the tailored drone and aerial footage or satellite imagery, not directly designed for the city users. Considering the abundance of already existing digital datasets dedicated for the city inhabitants and managers, it seems appropriate to try to make use of the available resources instead of putting effort into creating new ones.

FNN vs CNN context scraping: the framework

In the research on the possibilities of utilization of deep neural networks in spatial composition context processing [22], two different network architectures were evaluated by the author: FNNs and CNNs. Both models had a similar size (number of trainable parameters) and were trained for a comparable amount of time. The conducted preliminary tests indicated that CNNs could manage multidimensional spatial data much better than the FNNs (Fig. 1.). The CNNs showed greater endurance to overfitting [23] and would better generalize given previously unseen test examples. This observation seems to be in line with the findings of other research teams [7-9]. The discrepancy in the performance could be explained by the presence of the convolutional filters in place of the traditional, fully-connected layers. In simplified terms, the convolutional filters slide across the input data and learn to detect local features, instead of having one fully connected layer to process the whole input at once. The feature maps produced by the sliding kernels are then scanned by the next layer of convolutional filters, and so forth. The deeper the network is, the more complex spatial features can be processed by it [24].

The convolutional filters roaming across the scraped urban data could collect and process relevant local features in an analogous way to a city user roaming the city streets, with the difference being that a CNN agent would not process the real, 3D city space, but rather the multidimensional, abstract city space scraped from the digitized urban services, like Google Maps, GIS models or even open BIM databases. To account for the variable dimensionality and size of the input data, the proposed network framework would have to either use a batch size of one and a fully convolutional network architecture (as commonly applied in semantic segmentation tasks [25]) or use global pooling [26] before the subsequent fully-connected layers are applied. Under appropriately chosen tasks, the training pairs can be generated in a fully unsupervised manner [22].

Conclusions

As a study conducted by the author suggests, CNNs are better-suited models for spatial tasks than their fully connected counterparts. Due to their efficacy in the processing of multidimensional spatial data, convolutional neural networks are commonly used in gathering and processing of urban data. However, the majority of them relies on the costly, dedicated interfaces to collect the sensory inputs. Instead, a more sensible approach seems to be relying on data scraping of already existing, human dedicated urban databases. The proposed alternative approach sits in between the direct use of the available, machine-learning-compatible APIs and the cumbersome development of custom sensory systems or manually labeled datasets. The envisioned unsupervised learning framework could offer a more intuitive access to urban data, closer conceptually to the perspective of a real city user and distinct from reinforcement learning [27].

Bibliography

[1] Techopedia, Technology Dictionary, Data Scraping, online, access 10.2021: https://www.techopedia.com/definition/33132/data-scraping.

[2] Cambridge University Press & Assessment, Cambridge Advanced Learner's Dictionary & Thesaurus, Data Scraping, online, access 10.2021: https://dictionary.cambridge.org/pl/dictionary/english/data-scraping.

[3] Naumann H., Screen Scraper in Liu L., Ozsu M. T. Encyclopedia of Database Systems, Second Edition, Springer, 2018, 3325.

[4] Najork M., Web Crawler Architecture in Liu L., Ozsu M. T., Encyclopedia of Database Systems, Second Edition, Springer, 2018, 4608.

[5] Barracuda Networks Inc., Insights Into The Growing Number of Automated Attacks, Bot Attacks: Top Threats and Trends, barracuda.com, 2021, online, access 10.2021: https://assets.barracuda.com/assets/docs/dms/Bot_Attacks_report_vol1_EN.pdf.

[6] Souza T. G., Fonseca F. D. R., Oliveira Fernandes V., Pedrassoli J., Exploratory Spatial Analysis of Housing Prices Obtained from Web Scraping Technique, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B4-2021, 135-140.

[7] Fukushima K., Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position, Biological Cybernetics 1980, 36, 193-202.

[8] LeCun Y., Boser B., Denker J. S., et al., Backpropagation Applied to Handwritten Zip Code Recognition, Neural Computation 1989, 1(4), 541-551.

[9] Krizhevsky A., Sutskever I. and Hinton G. E., ImageNet Classification with Deep Convolutional Neural Networks, Advances in Neural Information Processing Systems 2012, 25(2), 84-90.

[10] Taoufiq S., Nagy B., Benedek C., HierarchyNet: Hierarchical CNN-Based Urban Building Classification, Remote Sens, 2020, 12(22), 3794.

[11] Pourazar H., Samadzadegan F., Javan F. D., A deep 2D/3D Feature-Level fusion for classification of UAV multispectral imagery in urban areas, Geocarto International, 2021.

[12] Egli S., Hopke M., CNN-Based Tree Species Classification Using High Resolution RGB Image Data from Automated UAV Observations, Remote Sens, 2020, 12(23), 3892.

[13] Khoshboresh-Masouleh M., Shah-Hosseini R., Building panoptic change segmentation with the use of uncertainty estimation in squeeze-and-attention CNN and remote sensing observations, International Journal of Remote Sensing, 2021, 42(20), 7798-7820.

[14] Qayyum A., Malik A., Saad N. M., Mazher M., Designing deep CNN models based on sparse coding for aerial imagery: a deep-features reduction approach, European Journal of Remote Sensing, 2019, 52:1, 221-239.

[15] Iino S., Ito R., Doi K. et al., CNN-based generation of high-accuracy urban distribution maps utilising SAR satellite imagery for short-term change monitoring, International Journal of Image and Data Fusion, 2018, 9(4), 302-318.

[16] Pan Y., Chen X., Sun Q., Zhang X., Monitoring Asphalt Pavement Aging and Damage Conditions from Low-Altitude UAV Imagery Based on a CNN Approach, Canadian Journal of Remote Sensing, 2021, 47(3), 432-449.
[17] Qi Y., Drolma S. C., Zhang X. et al., An investigation of the visual features of urban street vitality using a convolutional neural network, Geo-spatial Information Science, 23(4), 341 351.
[18] Law S., Seresinhe C. I., Shen Y., Gutierrez-Roig M., Street-Frontage-Net: urban image classification using deep convolutional neural networks, International Journal of Geographical Information Science, 2020, 34(4), 681-707.
[19] Ranjan N., Bhandari S., Zhao H. P., et al., City-Wide Traffic Congestion Prediction Based on CNN, LSTM and Transpose CNN, IEEE Access, 2020, 8, 81606-81620.
[20] Guo J., Liu Y., Yang Q. K., Wang Y., GPS-based citywide traffic congestion forecasting using CNN-RNN and C3D hybrid model, Transportmetrica A: Transport Science, 2021, 17(2), 190 211.
[21] Boulila W., Ghandorh H., Khan M. A. et al., A novel CNN-LSTM-based approach to predict urban expansion, Ecological Informatics, 2021, 64(101325).
[22] Dzieduszyński, T., Machine learning and complex compositional principles in architecture: Application of convolutional neural networks for generation of context-dependent spatial compositions, International Journal of Architectural Computing, manuscript accepted for publication, 2022.
[23] Strivastava N., Hinton G, Krizhevsky A., et al., Dropout: A simple Way to Prevent Neural Networks from Overfitting, Journal of Machine Learning Research 2014; 15, 1929-1958.
[24] Olah C., Mordvintsev A., Schubert L., Feature Visualisation, Distill, 2017, online, access: 10.2021: https://distill.pub/2017/feature-visualization/.
[25] Long J., Shelhamer E., Darrell T., Fully Convolutional Networks for Semantic Segmentation, IEEE CVPR, 2015, 1, 3431-3440.
[26] Lin M., Chen Q., Yan S., Network In Network, arXiv, v3, 2014.
[27] Kaelbling L. P., Littman M. L., Moore A. W., Reinforcement Learning: A Survey, Journal of Artificial Intelligence Research 4, 1996, 237-285.

DOI: 10.5604/01.3001.0015.7566

CORRECT QUOTATION FORMAT
Dzieduszyński Tomasz, 2022, Convolutional Neural Networks as Context-Scraping Tools in Architecture and Urban Planning, “Builder” 3 (296).
DOI: 10.5604/01.3001.0015.7566

Abstract: “Data scraping” is a term usually used in Web browsing to refer to the automated process of data extraction from websites or interfaces designed for human use. Currently, nearly two thirds of Net traffic are generated by bots rather than humans. Similarly, Deep Convolutional Neural Networks (CNNs) can be used as artificial agents scraping cities for relevant contexts. The convolutional filters, which distinguish CNNs from the Fully-connected Neural Networks (FNNs), make them very promising candidates for feature detection in the abundant and easily accessible smart-city data consisting of GIS and BIM models, as well as satellite imagery and sensory outputs. These new, convolutional city users could roam the abstract, digitized spaces of our cities to provide insight into the architectural and urban contexts relevant to design and management processes. This article presents the results of a query of the state-of-the-art applications of Convolutional Neural Networks as architectural “city scrapers” and proposes a new, experimental framework for utilization of CNNs in context scraping in urban scale.

Key words: Convolutional Neural Network, architecture, urban planning, smart city, data scraping, CAAD

Streszczenie: SPLOTOWE SIECI NEURONOWE JAKO NARZĘDZIA SŁUŻĄCE WYDOBYWANIU DANYCH ARCHITEKTONICZNO-URBANISTYCZNYCH. „Data scraping” to termin używany zazwyczaj w kontekście ruchu sieciowego, oznaczający proces automatycznej ekstrakcji danych ze stron internetowych i interfejsów, zaprojektowanych do stosowania przez człowieka. Obecnie blisko dwie trzecie ruchu internetowego jest generowanych przez boty, a nie przez ludzi. Na podobnej zasadzie głębokie splotowe sieci neuronowe (CNN) mogą być stosowane jako narzędzia wyszukujące w miastach stosowne konteksty urbanistyczne. Filtry splotowe, odróżniające CNN od sieci w pełni połączonych (FNN), sprawiają, że są one obiecującymi kandydatami do wykrywania cech ukrytych w zasobnych i łatwo dostępnych danych *smart city*, składających się z modeli GIS i BiM oraz obrazów satelitarnych oraz innych danych sensorycznych. Filtry splotowe mogą przemierzać abstrakcyjne, cyfrowe przestrzenie naszych miast, dostarczając kontekstów przydatnych w projektowaniu oraz zarządzaniu architektoniczno-urbanistycznym. Artykuł prezentuje wyniki kwerendy źródeł dotyczących najnowszych zastosowań splotowych sieci neuronowych w wydobywaniu danych miejskich i proponuje nowe, eksperymentalne ramy dla wykorzystania CNN w ekstrakcji kontekstów urbanistycznych.

Słowa kluczowe: splotowa sieć neuronowa, architektura, urbanistyka, smart city, wydobywanie danych, CAAD