AAU crenau
ambiances
architectures
urbanités

Masters internship - Deep Learning on the Place Pulse 2 dataset

Keywords: pairwise learning, visual urban perception, zero-shot learning, multi-task learning, computer vision, deep learning

Context and internship description:

In urban planning, the qualification of urban space is an important task that informs experts and local authorities on how public space is perceived by a pedestrian. This task involves measuring subjective qualities of urban space such as naturalness, spaciousness, tidiness, etc. A possible application is the identification of priority areas in the urban policy of the city.

This field is quickly developing thanks to the availability of images captured from the pedestrian point of view (street view imagery - SVI). For instance, services such as Google Street View (GSV) offer dense coverage of urban areas worldwide. Besides this, thanks to the high performance of convolutional neural networks (CNN) for image processing, many researchers have started to develop methods for automatic qualification of urban space.

To accompany an ongoing PhD thesis related to the study of urban space properties using machine learning tools, an in-depth analysis of a specific corpus is planned in 2022 as part of a Masters internship.

To date, the most comprehensive dataset in this field is Place Pulse 2.0 (PP2) [1]. It contains pairwise comparisons of images collected by crowdsourcing via an online platform. The user is presented with 2 GSV images, and must choose the one that corresponds the best to the proposed quality. The dataset contains 110,000 images and 1.1 million comparisons for the 6 following qualities: safer, depressing, boring, lively, wealthier, beautiful.

This dataset is generally used to train neural networks. The most common task is binary classification: the network takes 2 input images and must determine the winning image, for a given quality. Based on the high correlation between the 6 categories, some authors have focused on the use of multi-task learning [2,3], while others have proposed to use a zero-shot learning technique [4]. However, despite these efforts to use recent Deep Learning methods, current models offer rather low performances, around 70% accuracy.

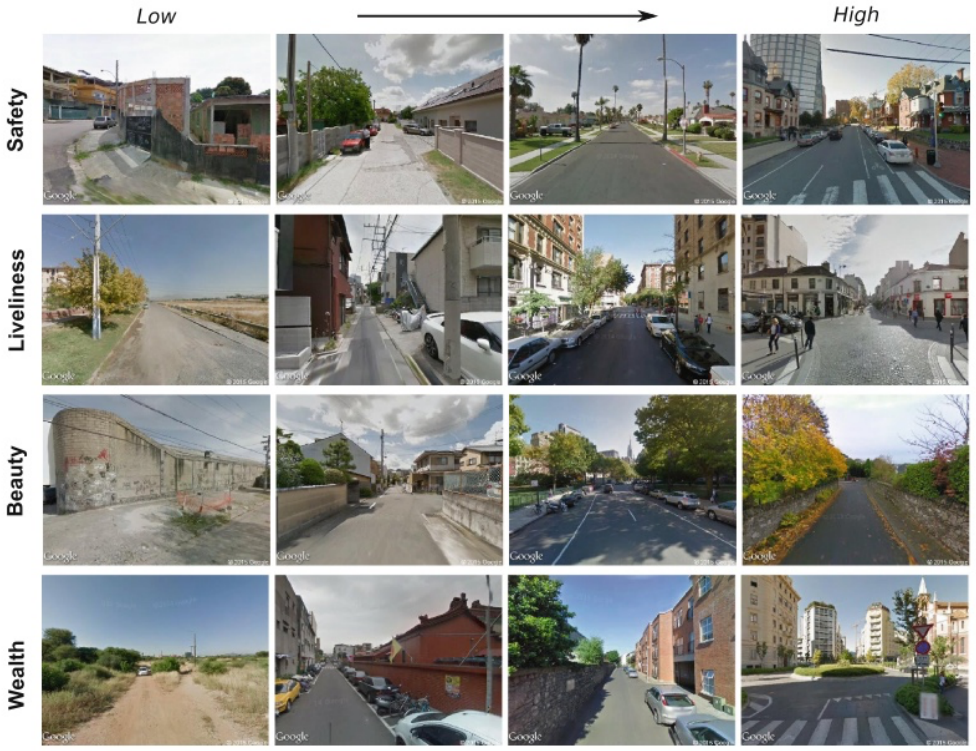

Figure 1: Images from the PP2 dataset, ordered by qualities from low to high according to the participants' ratings [1]

Responsibilities:

The goal of this internship is to explore different ways to work with this dataset and to propose a model with state of the art performances:

- Bibliography on deep learning methods for pairwise learning, existing work on PP2, the different architectures that could be used with this dataset and improve the state of the art results.
- Choose an architecture / method and apply it on the PP2 dataset.
- Compare the results obtained with those of the literature.

A publication of the results at the end of the internship is expected.

Qualifications:

- Strong skills in Python
- Some experience with a Deep Learning Framework (PyTorch preferred, or Tensorflow)
- A prior experience working on a computer vision project is a plus

About the internship:

- Start date: march or april 2022
- 5 to 6 months, full time
- Minimum legal internship gratification (about 550€/month)
- Location: AAU lab, Crenau team, école d'architecture de Nantes and remote work from 08/01 to 08/19.
- Supervisors: Benjamin Beaucamp, Vincent Tourre, Thomas Leduc, Myriam Servières

- Send your CV and cover letters to benjamin.beaucamp@ec-nantes.fr, vincent.tourre@ec-nantes.fr, thomas.leduc@crenau.archi.fr, myriam.servieres@ec-nantes.fr

References :

[1] Dubey, A., Naik, N., Parikh, D., Raskar, R., & Hidalgo, C. A. (2016). Deep learning the city: Quantifying urban perception at a global scale. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9905 LNCS, 196–212. https://doi.org/10.1007/978-3-319-46448-0_12

[2] Guan, W., Chen, Z., Feng, F., Liu, W., & Nie, L. (2021). Urban Perception: Sensing Cities via a Deep Interactive Multi-task Learning Framework. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 17(1s). https://doi.org/10.1145/3424115

[3] Min, W., Mei, S., Liu, L., Wang, Y., & Jiang, S. (2020). Multi-Task Deep Relative Attribute Learning for Visual Urban Perception. IEEE Transactions on Image Processing, 29, 657–669. https://doi.org/10.1109/TIP.2019.2932502

[4] Talebi, H., Amid, E., Milanfar, P., & Warmuth, M. K. (2020). Rank-Smoothed Pairwise Learning in Perceptual Quality Assessment. Proceedings - International Conference on Image Processing, ICIP, 2020-Octob, 3413–3417. https://doi.org/10.1109/ICIP40778.2020.9191231