Automatic production of environmental indicators from freely available remote sensing data: from a global to a local scale

Master internship

General information

- Keywords: Remote sensing, Deep learning, Sentinel 2, Local climate zones, Africa
- Duration of the internship: 6 months (standard stipend). To start between February and April 2021.
- Institutes: Université de Paris, Laboratoire d'Informatique Paris Descartes (LIPADE), équipe Systèmes Intelligents de Perception et Institut national d'études démographiques, unité Démographie des populations du Sud (Demosud)
- Location: 45 rue des Saints-Pères, 75006 Paris (LIPADE) and 9, cours des Humanités, 93322 Aubervilliers (INED)
- Supervision: Sylvain Lobry, Camille Kurtz, Laurent Wendling - (first.lastname@u-paris.fr), Géraldine Duthé, Valérie Golaz (first.lastname@ined.fr)
- Application: Please send a cover letter and a CV to stage-diip "at" listes.ined.fr. You will receive a confirmation by email. The position is opened until filled.

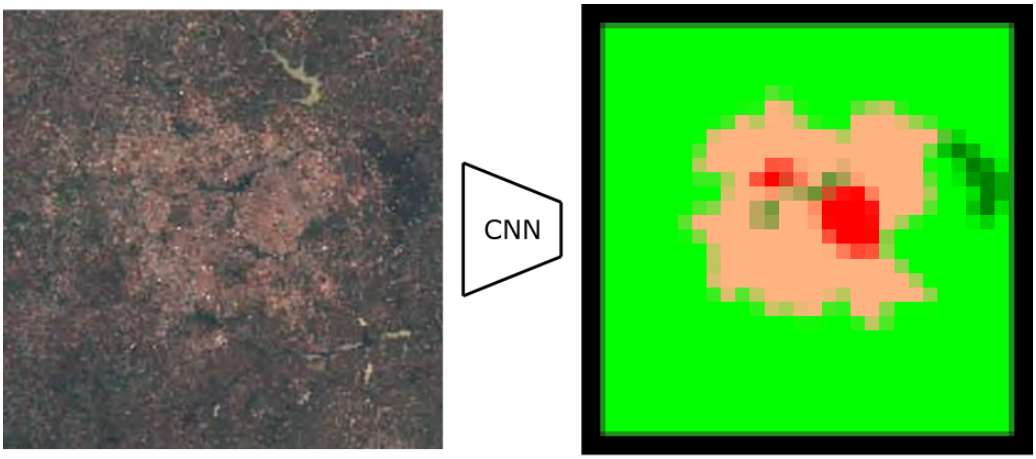
Proposed topic

Motivation

Recent sanitary crises and the on-going climate change process have major impacts on human populations worldwide. In Africa, the accessibility of demographic data is increasing. Population censuses are being performed regularly in most countries, at the national scale, every 10 years or so. Data can also be available at the local scale at a higher frequency, through demographic surveillance sites (HDSS). Local and national surveys add to this landscape. In order to understand the impact contextual and environmental change have on local populations, it is important to precisely quantify them. This collaborative project aims at studying the feasibility of automatically producing repeatable indicators from remote sensing data in Africa to allow for spatially complete and temporally up-to-date information. To this effect, we will use freely available Sentinel 2 images (produced by the European Space Agency) to produce standardized environmental indicators.

Background and state of the art

Several indicators and legends have been developed to help understanding geographical realities in a consistent (i.e. not location dependent) manner. Among them, local climate zones (LCZ) have been proposed by WUDAPT (World Urban Database and Access Portal Tools) to systematically label urban areas [1]. Their goal is to provide a map of the world following this legend, in open-access, that can later be used by researchers for a wide range of studies. This data has been used



(a) Sentinel-2 image of Ouagadougou (b) Sample of an LCZ classification

Figure 1: Sample of an LCZ classification based on a Sentinel-2 image obtained with a CNN (Contribution A of this internship).

to understand energy usage [2], climate [3] or geoscience modeling [4]. An important amount of work has been dedicated in the recent years to the automatic generation of such data, from sensors such as Landsat 8 or Sentinel 2. In a research competition organized by the IEEE IADF, several methods have been proposed to map LCZ from Landsat, Sentinel 2 and OpenStreetMap data [5]. Another recent study focused on the usage of Convolutional Neural Networks (CNNs) to tackle the task of automatically mapping LCZ using deep learning [6]. However, these works mostly focused on developed urban areas. For instance, the challenge of [5] targeted Berlin, Hong Kong, Paris, Rome, São Paulo, Amsterdam, Chicago, Madrid, and Xi'An. This is problematic, as developed cities are generally well mapped through governmental censuses, and that spatial generalization of machine learning based methods is a challenge [7]. It is therefore necessary to develop adapted methods for developing areas [8].

Work to be done

The work to be conducted during the proposed M2 internship will lead to the following three contributions:

- Contribution A: An automatic method for LCZ mapping for urban areas in developing countries, specifically targeted towards Africa with case studies on Ouagadougou, Burkina Faso and Antananarivo, Madagascar. An illustration of this first part is shown in Figure 1. This will be tackled through CNN models, based on the expertise developed in LIPADE.
- Contribution B: Reflection on the embedding of spatial relations, and temporal consistency of the result. In a second step, the student will have to study how to improve the results based on spatio-temporal relations. This could be done through spatial position descriptors [9] and through temporal networks, such as Long Short Term Memory networks.
- Contribution C: Reflection on the impact of LCZ on local inhabitants: a final contribution of this work will be to interpret the results through their impact for demographic studies. This will be done with the expertise of INED by including the indicators produced in demographic models. A graphical interface (GUI) will be implemented on top of the developed model that will allow to perform the LCZ classification given any remote sensing image.

Desired background

We are looking for a student in Master 2 or final year of MSc, or engineering school in computer science. The ideal candidate would have knowledge in image processing, computer vision, machine learning and Python programming and an interest in handling large amount of data, remote sensing and demography. An experience in statistical data analysis would be a plus.

Bibliography

[1] Benjamin Bechtel et al. "Mapping local climate zones for a worldwide database of the form and function of cities". In: *ISPRS International Journal of Geo-Information* 4.1 (2015), pp. 199–219.

[2] Paul John Alexander, Gerald Mills, and Rowan Fealy. "Using LCZ data to run an urban energy balance model". In: *Urban Climate* 13 (2015), pp. 14–37.

[3] Jan Geletič et al. "Spatial modelling of summer climate indices based on local climate zones: expected changes in the future climate of Brno, Czech Republic". In: *Climatic Change* 152.3-4 (2019), pp. 487–502.

[4] Hendrik Wouters et al. "The efficient urban canopy dependency parametrization (SURY) v1. 0 for atmospheric modelling: description and application with the COSMO-CLM model for a Belgian summer". In: *Geoscientific Model Development* 9.9 (2016), pp. 3027–3054.

[5] Naoto Yokoya et al. "Open data for global multimodal land use classification: Outcome of the 2017 IEEE GRSS Data Fusion Contest". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11.5 (2018), pp. 1363–1377.

[6] Chunping Qiu et al. "Feature importance analysis for local climate zone classification using a residual convolutional neural network with multi-source datasets". In: *Remote Sensing* 10.10 (2018), p. 1572.

[7] Emmanuel Maggiori et al. "Can semantic labeling methods generalize to any city? the inria aerial image labeling benchmark". In: *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. IEEE. 2017, pp. 3226–3229.

[8] John E Vargas-Muñoz et al. "Correcting rural building annotations in OpenStreetMap using convolutional neural networks". In: *ISPRS journal of photogrammetry and remote sensing* 147 (2019), pp. 283–293.

[9] Michaël Clément, Camille Kurtz, and Laurent Wendling. "Fuzzy directional enlacement landscapes for the evaluation of complex spatial relations". In: *Pattern Recognition* 101 (2020), p. 107185.