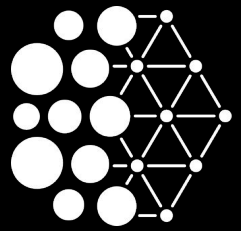
Institut
québécois
d'intelligence
artificielle



Mila

MAJOR UPDATE

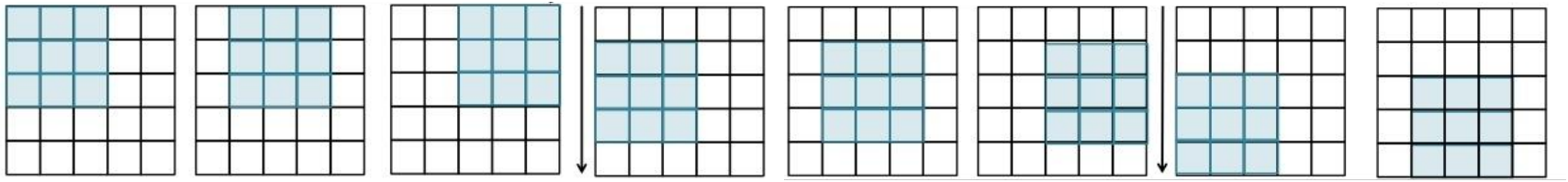
Horoma project
Block 2

Francis Grégoire
Mathieu Germain

Context

- While evaluating models, we found that we couldn't overfit on the training set (i.e. models were performing as well on the valid and test set).

- Why? Horoma used overlaps of 1 pixel (left→right, up→down) to generate 32 x 32 x 4 pixel patches (think of strides=(1, 1) in a CNN):



- It resulted in several pixel patches being extremely similar.

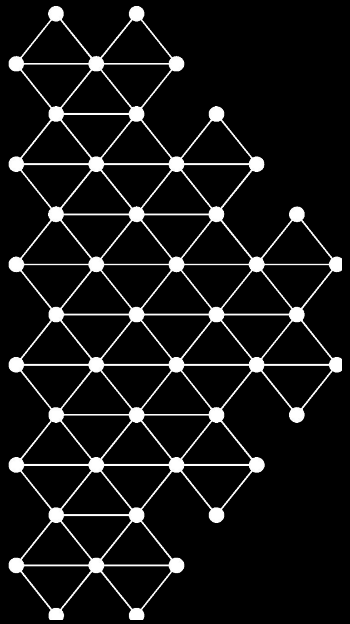
- Since training, valid and test sets were almost perfectly balanced, most pixel patches in the test set had almost identical pixel patches in the training set.

What's new?

- We created new (and simpler) datasets for you.
- Forget the confusing relative height/height in the inputs/outputs:
  - Inputs: each pixel in a pixel patch contains only RGB values (i.e. 32 x 32 x 3).
  - Outputs: each pixel patch has a tree specie (no more density and height).
- However, the size of the labeled data is dramatically reduced:
  - 480 or 1331 (with ~50% overlaps) examples in valid set.
  - 498 examples in test set.
- The new objective is to predict the tree specie of a given 32 x 32 pixel patch.

Institut
québécois d'intelligence
artificielle

Mila



Updated slides

Data (pixel)

- Each pixel of an image has **3** ~~4~~ values associated with it:
  - **RGB colors** (3 values). Those values are in [0, 255].
  - ~~**Height** (1 value). The height values were obtained using photogrammetry and were georeferenced; they are measured w.r.t. the sea level.~~

Data (pixel patches)

- 32 x 32 pixel patches were extracted from labeled image subsections.
- **Inputs**: each pixel in a 32 x 32 pixel patch has **3** ~~4~~ dimensions.
- **Outputs**: each 32 x 32 pixel patch has **1** ~~3~~ label:
  - Tree species.
  - ~~Tree density.~~
  - ~~Tree height w.r.t. the forest floor.~~

Data (input format)

- Inputs are provided as binary numpy.memmap files in **float32**.
- Memory-mapped files are used for accessing small segments of large files on disk, without reading the entire file into memory:
  - **train_x.dat**: **150,900** x 32 x 32 x **3** ~~1,614,216 x 32 x 32 x 4~~.
  - **train_overlapped_x.dat**: 544,749 x 32 x 32 x 3 (pixel patches with 50% overlap).
  - **valid_x.dat**: **480** x 32 x 32 x **3** ~~201,778 x 32 x 32 x 4~~.
  - **valid_overlapped_x.dat**: 1331 x 32 x 32 x 3 (pixel patches with 50% overlap).
- You also have access to files containing the image and the pixel subregion where each pixel patch has been extracted:
  - **train_regions_id.txt** & **train_overlapped_regions_id.txt**.
  - **valid_regions_id.txt** & **valid_overlapped_regions_id.txt**.

Data (output format)

- Outputs are provided as **2** ~~3~~ text files (can be easily read from a terminal).
- **valid_y.txt**: contains 480 tree species (2 characters).
- **valid_overlapped_y.txt**: contains 1331 tree species (2 characters).
- The *i*-th value in **valid_y.txt** and **valid_overlapped_y.txt** is associated to the *i*-th pixel patch in **valid_x.dat** and **valid_overlapped_x.dat**.