Large image time series analysis for updating vineyard geographic database

Work supported by diiP, IdEx Université de Paris, ANR-18-IDEX-0001

Authors

Lucas Colomines[1]

Supervisors :
Camille Kurtz[1]
Nicole Vincent[1]
Michele Linardi[1]

Collaboration :
Anne Puissant[2]

[1] **Université de Paris, LIPADE** - {first name.last name}@u-paris.fr

[2] **Université de Strasbourg, LIVE** - {first name.last name}@live-cnrs.unistra.fr

Introduction

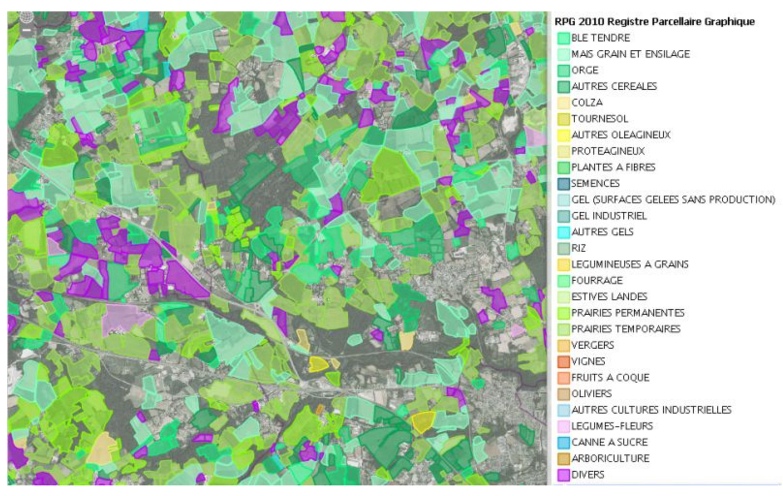
Objectives :
- Segment vineyards from satellite image time series
- Use both spatial and temporal information for segmentation
- Correct the “Registre Parcellaire Graphique” (a.k.a R.P.G.)

Constraints :
- There is a lack of annotated data. It is a problem for training and evaluation of deep learning models

Data

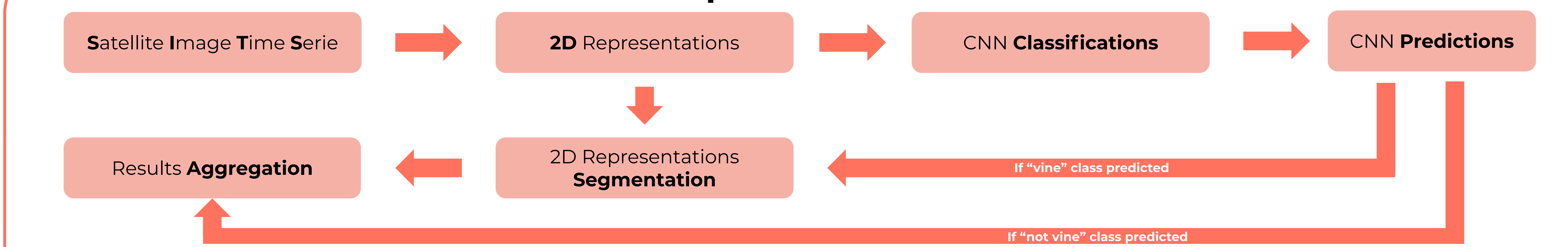
R.P.G. :
- Geographic database containing agricultural plots represented with polygons
- Updated with farmers declarations : not accurate

Satellite Image Time Series (SITS) :
- From Venµs satellite
- Spatial resolution : 1px = 5m
- Twelve spectral bands
- Acquisition : One image every two days



RPG

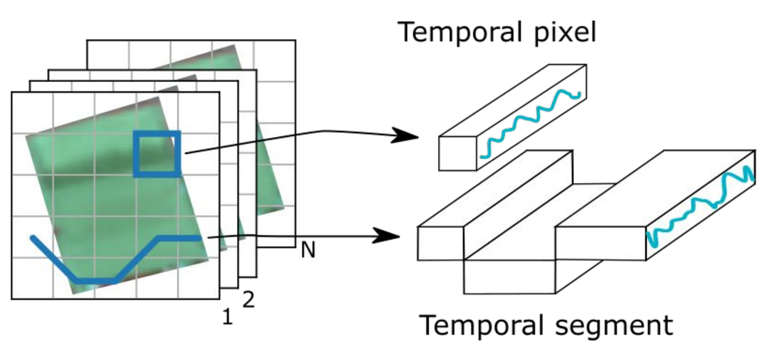
Proposed Method

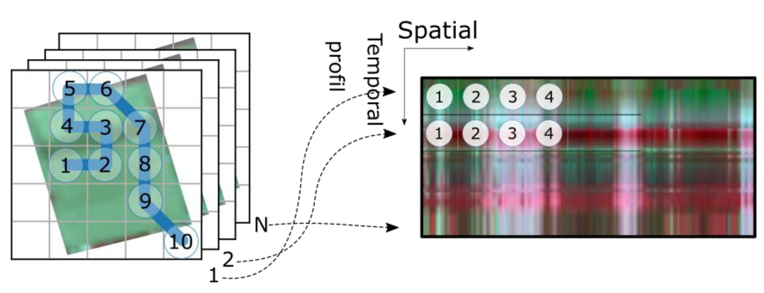


2D Representations

Why 2D Representations :
- 3D CNNs need a lot of datas
- Generated from a random “segment” -> “data augmentation”

How to build 2D representations :
- Find a serie of pixels by using a random walk algorithm (in the spatial domain)
- Plot the serie of pixels on the horizontal axis and the temporal values of each pixels on the vertical axis





CNN Classification & Results Aggregation

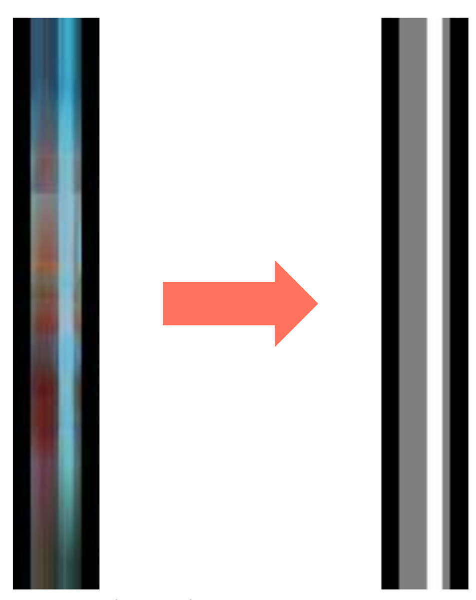
CNN Classification :
- Pre-Trained Squeezenet
- Classify 2D representations
- ~90% classification accuracy
- Outputs a “confidence score”

Results aggregation :
- After computing a prediction for each 2D representations
- Compute the average “vine” confidence score for every 2D representations passing on the pixel
- Repeat for each pixel
- Leads to a binary semantic segmentation result

2D Representations Segmentation

What happens if a segment is generated between vine and “not vine” ?
Part of the pixels on the segment will be misclassified.

Solution :
- If the 2D representation is classified as “Vine”
- Segment the different crops in the 2D representation using the average color per columns
- Consider the largest one as “Vine”



2D representation vine crossed by a road | Segmentation Gray : Vine | White : road

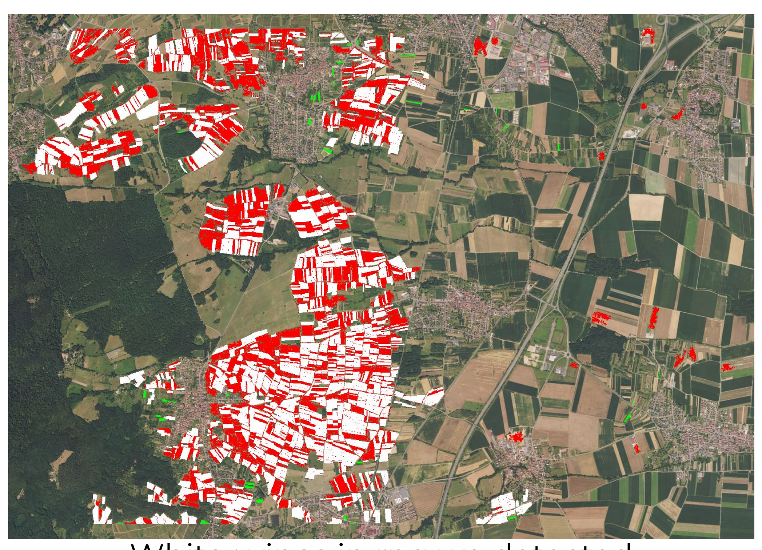
Evaluation & Results

Once the results are aggregated, we find the binarization threshold maximizing the accuracy on a validation area.

We manually built a ground truth. Then we compare the R.P.G. and our segmentation to this ground truth to evaluate it.

Even though our precision is a little below the rpg, we detect more vines and have better global accuracy.

Precision : 0.87 (us) | 0.94 (R.P.G.)
Recall : 0.89 (us) | 0.54 (R.P.G.)
Accuracy : 0.97 (us) | 0.93 (R.P.G.)



White : vines in rpg we detected
Red : Vines missing in the RPG we detect
Green : Vines we missed

Reference

Chelali, M., Kurtz, C., Puissant, A. & Vincent, N. (2021) **Deep-STaR: Classification of image time series based on spatio-temporal representations, Computer Vision and Image Understanding**, 208--209:103221