



HAL
open science

Development of a software based on automatic multi-temporal aerial images classification to assess retrospective environmental exposures to pesticides in epidemiological studies

Élodie Faure, Rémi Ratajczak, Carlos F Crispim-Junior, Olivia Pérol, Laure Tougne, Béatrice Fervers

► To cite this version:

Élodie Faure, Rémi Ratajczak, Carlos F Crispim-Junior, Olivia Pérol, Laure Tougne, et al.. Development of a software based on automatic multi-temporal aerial images classification to assess retrospective environmental exposures to pesticides in epidemiological studies. CLARA 2018 Cancer Research Forum, Apr 2018, Lyon, France. hal-01784654

HAL Id: hal-01784654

<https://hal.science/hal-01784654>

Submitted on 3 May 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Development of a software based on automatic multi-temporal aerial images classification to assess retrospective environmental exposures to pesticides in epidemiological studies

E. Faure¹, R. Ratajczak^{1,2,3}, C. Crispim-Junior², O. Perol¹, L. Tougne² and B. Fervers¹

¹Cancer and Environment Department, Centre Léon Bérard, Lyon, France

²Univ Lyon, Lyon 2, LIRIS, F-69676 Lyon, France

³French Environment & Energy Management Agency, Angers, France

1. Context

- Pesticides have been suggested as risk factors for different pathologies (cancers, infertility, Parkinson's...)
- Geographical Information Systems (GIS) are increasingly used to assess environmental exposure to agricultural pesticides (EEAP) in epidemiological studies by taking into account long latency period
- There is a lack of annotated land cover data in France before 1990
- Large scale retrospective epidemiological studies like TESTIS (Béranger et al. 2014) are carried out over 45 years in the past and involve hundreds of subjects (1500 for TESTIS)
- The manual annotation of a single historical aerial image (National Geographic Institute – IGN) representing a 9 km² area takes 6 to 10 hours (1 subject/1 date/1 location in TESTIS)

Goal: Assess the lack of annotated historical land cover data in France by developing an automated image processing software for fast annotations of historical aerial images

2. Method - Creation of a new dataset : HistAerial

- ✓ 4.9 millions annotated land cover patches issued from historical aerial images acquired between 1970 and 1990 by the IGN (non-overlapping, monochromatic)
- ✓ 7 classes (Arable, Water, Forest, Grassland, Urban, Orchard, Vineyard)
- ✓ 3 scales (100 x 100 pixels, 50 x 50 pixels, 25 x 25 pixels)
- ✓ Patches represent inhomogeneous textures

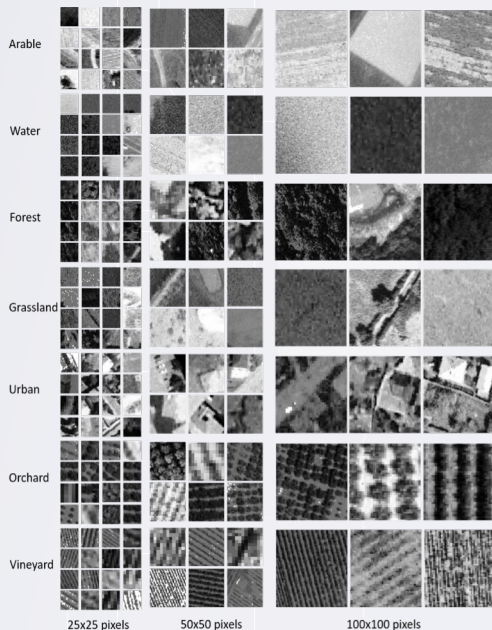


Fig. 1 : Examples of historical land cover patches in the HistAerial dataset

3. Method - An extensive comparative study : Overview

- ✓ 12 state-of-the-art local texture filters, low dimensional representations based on the Local Binary Pattern (LBP), and one variant (LBP_riu2)
- ✓ 2 novel local texture filters, namely LCOLBP and R-CRLBP
- ✓ 4 classifiers (K-Nearest Neighbors, Support Vector Machine, Multi-Layer Perceptron, Random Forest), trained with grid search
- ✓ 3 finetuned end-to-end deep convolutional neural networks
- ✓ Evaluation on an equilibrated and randomly sampled subset of HistAerial (6 000 images / class)

4. Results - An extensive comparative study : Evaluation

- ✓ Our understandable handcrafted filter LCOLBP achieves similar results with deep learning method AlexNet
- ✓ LCOLBP generates features vector 17 times shorter than AlexNet, thus yielding to faster training and inference

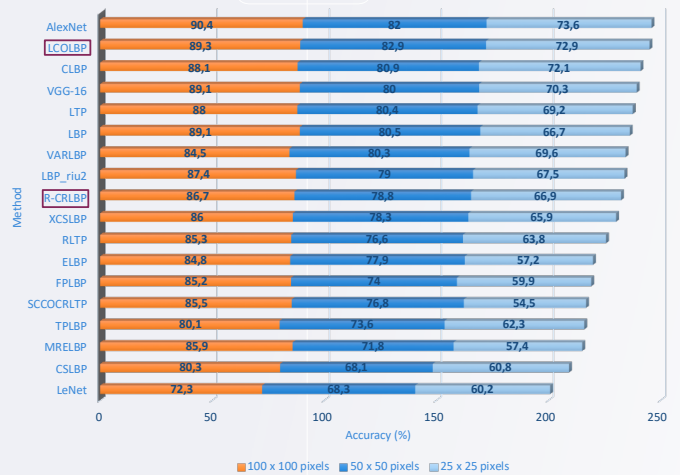


Fig. 2 : Best accuracy rates obtained on the HistAerial dataset

5. Results - Software

- ✓ Development of a scribble-patch based software
- ✓ Integrate best evaluated methods
- ✓ Enable annotation of 9 km² in 5 to 10 minutes instead of 6 to 10 hours on a single CPU thread (1.7Ghz) ; scribble, training and test times included

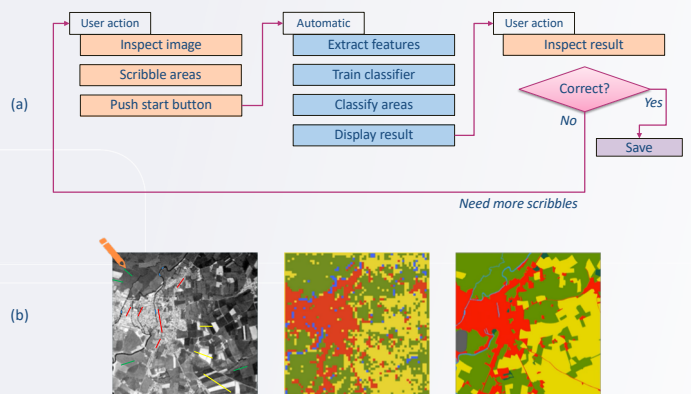


Fig. 3 : (a) Software workflow, (b) Left - scribbled historical aerial image ; Center - result with the software learning solely from scribbles ; Right - ground truth, unseen by the software

6. Conclusion and perspectives

- ✓ Novel large scale dataset for historical aerial images classification
- ✓ Extensive comparative study including two novel texture filters
- ✓ Innovative software prototype for fast annotations of historical aerial images
- Combine deep learning methods with handcrafted methods
- Integrate results in a Geographical Information System
- Creation of new land use datasets
- Use the software on large scale retrospective epidemiological studies