



HAL
open science

Automatic Land Cover Reconstruction From Historical Aerial Images: An Evaluation of Features Extraction and Classification Algorithms

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

► **To cite this version:**

Rémi Ratajczak, Carlos F Crispim-Junior, Élodie Faure, Béatrice Fervers, Laure Tougne. Automatic Land Cover Reconstruction From Historical Aerial Images: An Evaluation of Features Extraction and Classification Algorithms. IEEE Transactions on Image Processing, In press, 10.1109/TIP.2019.2896492 . hal-02003932

HAL Id: hal-02003932

<https://hal.science/hal-02003932>

Submitted on 7 May 2019

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.

Automatic Land Cover Reconstruction From Historical Aerial Images: An Evaluation of Features Extraction and Classification Algorithms

Rémi Ratajczak, Carlos Fernando Crispim-Junior, Elodie Faure, Béatrice Fervers, Laure Tougne

Abstract—The land cover reconstruction from monochromatic historical aerial images is a challenging task that has recently known an increasing interest from the scientific community with the proliferation of large scale epidemiological studies involving retrospective analysis of spatial pattern. However, the efforts engaged by the computer vision community in remote sensing applications are mostly focused on prospective approaches through the analysis of high resolution multi-spectral data acquired by advanced spatial programs. Hence, four contributions are proposed in this article. They aim at providing a comparison basis for the future development of computer vision algorithms applied to the automation of the land cover reconstruction from monochromatic historical aerial images. Firstly, a new multi-scale multi-date dataset composed of 4.9 million non-overlapping annotated patches of the France territory between 1970 and 1990 has been created with the help of Geography experts. This dataset has been named HistAerial. Secondly, an extensive comparison study of state-of-the-art texture features extraction and classification algorithms including deep convolutional neural networks (DCNNs) has been performed. It is presented in the form of an evaluation. Thirdly, a novel low-dimensional local texture filter named Rotated-CorneR Local Binary Pattern (R-CRLBP) is presented as a simplification of the Binary Gradient Contours filter through the use of an orthogonal combination representation. Finally, a novel combination of low-dimensional texture descriptors, including the R-CRLBP filter, is introduced as a Light Combination of Local Binary Patterns (LCoLBP). The LCoLBP filter achieved state-of-the-art results on the HistAerial dataset while conserving a relatively low-dimensional feature vector space compared with the DCNN approaches (17 times shorter).

Index Terms—Features extraction, Texture filters, Deep Convolutional Neural Networks, Deep Learning, Machine Learning, Land Cover, Historical Aerial Images

I. INTRODUCTION

The automatic reconstruction of the land cover is a process at the boundary between geography and computer sciences. It aims to provide a reusable solution to the increasing demand of precise data used to monitor the earth’s surface evolution. Considered as a challenging problem, it has known a huge interest in the last two decades due a large amount of publicly

R. Ratajczak is with Univ Lyon, Lyon 2, LIRIS, F-69676 Lyon, France. He is also with the Leon Bérard Center, France, and with the French Environment and Energy Management Agency, 20, avenue du Grésillé- BP 90406 49004 Angers Cedex 01 France (email: remi.ratajczak@liris.cnrs.fr).

C.F. Crispim-Junior is with Univ Lyon, Lyon 2, LIRIS, F-69676 Lyon, France (email: carlos.crispim-junior@liris.cnrs.fr).

E. Faure is with the Leon Bérard Center, France (elodie.faure@lyon.unicancer.fr).

B. Fervers is with the Leon Bérard Center, France. She is also with the Lyon 1 University and the Cancer Research Center of Lyon, UMR 1052 Inserm 5286 CNRS, France (beatrice.fervers@lyon.unicancer.fr).

L. Tougne is with Univ Lyon, Lyon 2, LIRIS, F-69676 Lyon, France (email: laure.tougne@liris.cnrs.fr).

available datasets acquired through advanced spatial programs like LandSat¹, QuickBird² and Sentinel³.

These spatial programs have demonstrated the efficiency of the satellites to provide a large quantity of relatively high resolution data. For instance, the spatial resolution of the satellite Landsat-8 is 30 meters, 10 meters for the satellite Sentinel 2-B and up to 1.5 meters for SPOT-6⁴. The satellites are embedded with multi-spectral acquisition systems (*e.g.* colors, infrared, radar, *etc.*) and ego-sensors like GPS and accelerometers. They enable the observation of the Earth in a repeatable manner through the acquisition of geolocalized orthorectified temporal series (*e.g.* 16 days between two images of the same area for Landsat-8). Researchers have been able to demonstrate successful segmentation and classification results at both coarse and fine grained scales (*e.g.* individual buildings [4][16], pixel [13]) by wisely combining the features obtained from the multiple sources of images available. Recently, Kussul et al. [32] proposed a deep learning based approach to classify crop fields from land cover data. Albert et al. [2] used a deep convolutional neural networks (DCNN) approach for the analysis of the urban environment based on satellite imagery. Slimene et al. [54] proposed an active learning approach to segment cultivated parcels from satellite images based on the Normalized Difference Vegetation Indexes (NDVI), multi-spectral standard deviation and mean. Wegner et al. [60] presented an insight on recent computer vision algorithms applied to high resolution remote sensing images, including landmark localization [49] and hybrid machine learning approaches [64].

Opposed to the multi-spectral data, the monochromatic historical aerial images captured before 1990 may show a lack of discriminative power. Most of them are panchromatic and are represented with only one intensity channel. The resolution of these scanned images complicates their analysis at a fine grained scale (see Figure 1) compared to the current satellite and aerial images. The recent algorithms that have been designed to work on data captured with multi-spectral sensors cannot be used as is with the historical aerial images because they expect a multi-spectral input. However, historical aerial images are gaining importance in the context of retrospective studies carried out in fields like environment, health and urbanism. Until now, epidemiological studies mostly used land cover data from satellites. As the relation between environmental exposures to pesticides and pathologies such as cancer are complex due to long latency, the characterization of longer

¹<https://landsat.usgs.gov/>

²<http://www.geoimage.com.au/satellite/quickbird>

³<https://sentinel3.cnes.fr/>

⁴<http://www.intelligence-airbusds.com/en/147-spot-6-7-satellite-imagery>

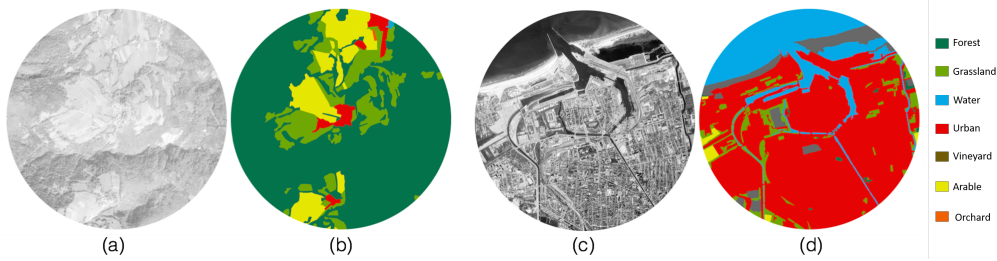


Fig. 1: Two examples of historical aerial images (a)(c) and their respective land cover ground truths (b)(d).

term retrospective exposures is essential to understand the causes of the cancer. Large scale studies, like TESTEPERA [7] and TESTIS [8], have already highlighted the need to reconstruct the land cover from historical aerial images through the use of Geographical Information Systems (GIS). Other studies used a combination of historical and actual land cover images for the analysis of the impact of the pesticides on the Parkinson disease [12]: The original historical land cover they used did not contain enough classes (*i.e.* 3 classes) to thoroughly investigate exposure to pesticides in their study. To accomplish this task, Brouwer et al. used images that were already annotated with 9 classes for the recent data and 3 classes for the historical data. They developed a statistical model to refine the historical land cover of the Netherlands with the more recent and complete annotations issued from the manual analysis of modern images.

The manual segmentation of one historical aerial image representing a 9 square kilometers area remains a tedious task which generally takes 6 to 10 hours to be carried out by a geographer. Due to the heavy workload and the lack of open data before 1990, it becomes important to study, propose and develop novel algorithmic methods dedicated to the automation of the reconstruction of the land cover from monochromatic historical aerial images. Aside from the expected speed gain, the development of such methods may lead to more homogeneous results throughout time than a single manual interpretation, which may be prone to an interpretation bias, especially when the number of classes increases.

This paper proposes four contributions to improve the analysis of historical aerial data and motivate future studies and developments. Firstly, it presents a new challenging multi-date and multi-scale dataset annotated by Geography experts (see section II). It has been created to assess the lack of annotated land cover data for the analysis of historical aerial images. It is composed of 4.9 million Earth surface patches acquired over the French territory between 1970 and 1990. Secondly, it presents a cross comparison study of state of the art handcrafted and learned feature extraction and classification algorithms. Thirdly, a novel filter called Rotated-CorneR Local Binary Pattern (R-CRLBP) is proposed as a complementary features extraction method to the existing state of the art. The efficiency of the R-CRLBP filter has been assessed on the proposed dataset. Finally, the complementarity of the R-CRLBP filter with the existing filters has been assessed through an effective combination of texture features named

Light Combination of Local Binary Patterns (LCoLBP), also introduced in this paper.

In this study, the problem of accurate land cover assessment has been expressed as a classification problem to evaluate the capacity of various processing chains to recover the empirical information needed by the geographers. The French territory has been selected because it is supposed to be a quite challenging territory from a remote sensing point of view: Both natural and urban areas are neither symmetric nor have representative and repetitive shapes compared with the ones analyzed by Yan et al. [63]. Moreover, during the studied period, France has known demographical modifications due to economic factors that have eventually changed the country shape on both urban and agricultural environments.

The remainder of this paper is organized as follows. Section II describes the proposed dataset. Section III introduces the compared algorithms. Section IV discusses the results. Section V summarizes and proposes future work perspectives.

II. THE DATASET: HISTAERIAL

This section presents the HistAerial dataset. It has been specifically created for this study. It is fully available for research purposes⁵⁶.

A. Image retrieval

The HistAerial dataset has been created using monochromatic historical aerial images of the French territory. These images were acquired between 1970 and 1990 with an optical camera mounted on an aircraft. They are now freely available on the french National Geographic Institute website (IGN)⁷. No land cover ground truth is provided for these images. Only few images are paired with infrared data even though the technology existed during the studied period. In consequence, no infrared images have been included in the proposed dataset.

Once retrieved, the historical aerial images have been projected in the RGF93 coordinates system [21], which is a widely used geographical reference system for the French territory.

B. Image properties

The images used to create the HistAerial dataset have the following properties: they are monochromatic; they have been

⁵<http://eidolon.univ-lyon2.fr/~remi1/HistAerialDataset/>

⁶<https://bit.ly/2QhueOe>

⁷<https://remonterletemps.ign.fr>

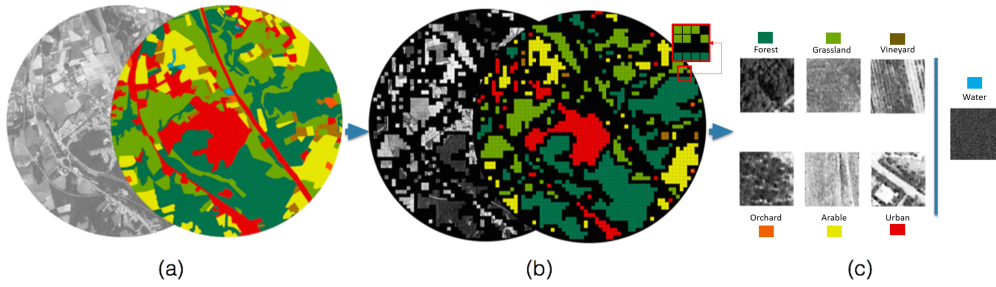


Fig. 2: The extraction process of patches: (a) original images; (b) visualization of corresponding patches where black squares represent excluded patches; (c) extracted patch samples.

captured when the sun was at the highest point during sunny periods in order to limit shadows and avoid possibly disturbing clouds; they are geolocated; for a given geographical coordinate, multiple images may have been acquired during time with up to 4 years between two acquisitions; the exact quality of the images is supposed unknown due to the possible modifications of the acquisition systems through the studied period (*e.g.* hardware updates) coupled with the uncontrollable outdoor conditions (*e.g.* wind, dust, *etc.*); and they have been acquired far in the past avoiding the possibility to acquire new images for the studied period.

These properties induce high intra-class and low inter-class variabilities in both space and time. It is not possible to classify the observed earth surfaces based on the most discriminative colors, Normalized Difference Vegetation Indexes (NDVI) [54] or joint multi-spectral distributions [27] because only one-color channel is available. The high temporal distance between two acquisitions of a same geographical area associated with the possible landscape modifications over the time tend to complicate the use of time series analysis in order to produce time-robust results like Kussul et al. [31] were able to obtain.

C. The patches

Geography experts from the department Cancer and Environment of the Léon Bérard Center have been asked to manually crop, segment and annotate historical aerial images using a 1.5 kilometer radius perimeter with 7 representative classes: Orchard, Arable, Grassland, Vineyard, Urban, Forest and Water (see Table I). A total number of 81 historical aerial images have been manually annotated in very high resolution (VHR). This number includes 56 densely annotated images with all the possible classes (when present), 15 partially annotated images for the Orchard class and 10 partially annotated images for the Vineyard class. Both the Orchard and the Vineyard classes were highly under represented on the first 56 images. Two historical aerial images with their manually segmented ground truths are presented on the Figure 1.

Then, inspired by the approaches based on image patches proposed by Barbier et al. [5] and Gonzalo et al. [22], patches of three arbitrarily chosen square sizes (*i.e.* 25 pixels \times 25 pixels; 50 pixels \times 50 pixels; 100 pixels \times 100 pixels) have been automatically and independently extracted in a supervised manner from the same annotated very high resolution images (see Figure 2). The extraction considered non-overlapping

patches (*i.e.* stride equals square size) corresponding to a single class using the following rule: if all the pixels of a patch - and not only the central pixel - belong to the same class, then store the patch, else discard the patch. Patches have been stored according to both their sizes and labels (see Table I). The non-linear factor between the number of patches for each patch size visible on Table I is a consequence of our choice to discard the patches representing more than one class with no overlapping between the patches of a same size. As visible on Figure 2, the number of patches obtained with this process of extraction is fully dependent of the non-linear semantic class frontiers represented on the ground truth images. Smaller patches tend to be overrepresented compared to bigger patches because it is easier for bigger patches to overlap multiple classes and thus being discarded. A factor 4 between the number of patches for each patch size could have been expected otherwise.

Since the number of patches per class and per size were not balanced, two subsets of the HistAerial dataset have been created (see Table II). These subsets could be considered as two independent datasets issued from the HistAerial dataset. The first subset is the size-balanced subset (see Table II). It was obtained following a random sampling strategy on the patches of each class and of each size based on the lowest number of patches available overall the HistAerial dataset (*i.e.* Water in 100 pixels \times 100 pixels). This subset aims to consider the same number (not proportion) of data

TABLE I: The complete HistAerial dataset.

Class	Number of patches per size (in pixels)		
	25 \times 25	50 \times 50	100 \times 100
Orchard	319 804	76 866	17 888
Arable	631 015	145 097	30 754
Grassland	348 349	71 334	11 984
Vineyard	174 288	40 528	8 889
Urban	891 500	204 746	43 254
Forest	443 760	95 945	18 554
Water	121 294	28 173	6 207
Total	2 930 010	662 689	137 530

TABLE II: The size-balanced and the class-balanced subsets of the HistAerial dataset.

Class	Number of patches per size (in pixels)			
	size-balanced	class-balanced		
		all sizes	25 \times 25	50 \times 50
Orchard	6 000	120 000	28 000	6 000
Arable	6 000	120 000	28 000	6 000
Grassland	6 000	120 000	28 000	6 000
Vineyard	6 000	120 000	28 000	6 000
Urban	6 000	120 000	28 000	6 000
Forest	6 000	120 000	28 000	6 000
Water	6 000	120 000	28 000	6 000
Total	42 000	840 000	196 000	42 000

for each size and for each class so that the lack of data for a given size compared to another would not influence the comparison of the filters and classifiers (see section IV). However, this approach alone does not consider the importance of the representation variabilities between the patches acquired for each size. Randomly sampling 6 000 patches out of 43 000 patches (100 pixels \times 100 pixels, class urban) is expected to induce a lower variability in the sampled data than randomly sampling 6 000 out of 891 000 patches (25 pixels \times 25 pixels, class urban). In order to overcome this issue, the class-balanced subset has been created (see Table II). For each size, patches have been randomly sampled based on the lowest number of patches available. The class-balanced subset is expected to represent the same yet approximated proportion (not number) of patches for each patch size. Note that these subsets have been sampled once for all so that the experiments discussed in Section IV) have all been performed on the same data.

The use of patches has several benefits. First, all the images (*i.e.* the patches) have the same size. No resizing operation that would not conserve the size ratio of the images is required to compare them. If any resizing operation would be required anyway, it would be performed on square images of equal size. This process should not modify the relative texture and shape properties from one patch to another. Second, the patches from the HistAerial dataset consist in non-overlapping images. They could be considered as spatially independent images that should capture the inter-class and intra-class variabilities without incorporating an explicit correlation between two patches (see Figure 3) in a similar manner to the dataset proposed by Porebski [47]. Third, patches allow simple but efficient multi-scale studies: Only the size of the patches has to be changed during the extraction step to acquire a new size of patches.

The downsides of this approach are summarized as follows. First, the number of patches per class depends on fixed sized annotated images, which are based on real-world data that cannot represent naturally each class in an equilibrated manner. The size-balanced and the class-balanced subsets of the HistAerial dataset aim to solve this issue. Second, since only the patches representing a unique class are retrieved, the larger the patches, the lesser patches are obtained (see Table I). In other words, the more the spatial context is captured in a patch, the lesser data could be obtained from an image. In practice, the number of patches decreases as their size increase (see Table I). That is the reason why there is no patches larger than 100 pixels \times 100 pixels in the proposed HistAerial dataset: Larger patches would yield to a very small dataset. The use of overlapping patches could eventually solve these issues. However, for the sake of a fair comparison basis not involving spatial redundancies between patches, it has been decided to keep the patches spatially independent like in [47]. Third, the land use land cover (LULC) classes may have specific features that can be captured at different spatial scales but may not be directly represented in a single patch. In order to evaluate this property, the features extraction step presented in IV-A is carried out using multi-scale handcrafted filters. Results show that the normalized feature vector and learned filters

from deep convolutional neural networks seem to implicitly handle multi-scale representations. Nonetheless, while square patches have obvious advantages to compare low level features extraction and classification algorithms, they may ignore the hierarchical nature of the LULC classes compared to object-based approaches, resulting in less accurate representations.

III. COMPARED ALGORITHMS

This section presents the feature extraction and classification algorithms used on the HistAerial dataset.

A. Handcrafted filters of the literature

The handcrafted filters that have been investigated in this study have been mainly introduced in the context of texture classification tasks. The use of texture-based filters has already been appreciated in previous work on remote sensing images. It was one of the earliest and most intuitive research track to automatize the analysis of these images [56]. Aerial images are indeed representing large scale areas made of spatially close objects observed from a high and generally almost perpendicular observation point. From this viewpoint, the earth surface is represented with specific and almost repetitive structural patterns, which implicitly correspond to the definition of the inhomogeneous textures in computer vision. Based on these observations, we considered the land use land cover (LULC) classes of the HistAerial dataset as regular texture classes. In result, a total of 14 state-of-the-art handcrafted texture filters based on the Local Binary Pattern (LBP) [41] have been compared on the HistAerial dataset. Similarly to the Textural Ordination based on Fourier Spectral Decomposition proposed by Coueron et al. [15] to analyze panchromatic images, the LBP-like filters have the property to summarize the low and high frequencies in an image with efficient non-linear directional operations through the spatial coordinates. The LBP-like filters used in this study are detailed below with the assertion that the input images are grayscale images. Other analogous and more classical filters such as the Gray Level Co-Occurrence Matrix (GLCM) [24] and the Gabor Wavelets [39] have not been included in this study in order to keep the parameters space consistent across the experiments. Moreover, previous studies have already assessed the effectiveness of these filters on remote sensing images [18] [1] [27] as well as on common texture datasets [19]. They demonstrated, overall, that the LBP-like filters tend to provide high accuracy rates on texture images making them good candidates to represent the LULC classes.

1) *Local Binary Pattern [41]*: The LBP filter aims to represent the non-directional local relations of a central pixel g_c to its P neighbors g_p that are evenly spaced on a circle of radius R . The relation between g_c and its neighbors is represented with a binary code made of P bits, where each bit represents the sign of the difference between g_c and g_p as presented in Equation (1). This relation is also called first order circular derivative [57]. The circular neighborhood is always read in the same order to generate a binary code in the original LBP [41]. P and R are defined as independent positive values. P is generally chosen as a multiple of 8 and R

with a strictly positive value. Tuning these values allow multi-scale analysis. Common paired values are $P = (8, 16, 24)$ and $R = (1, 2, 3)$. The neighborhood of the LBP filter could be represented with discrete coordinates (*i.e.* a neighbor is a pixel) or continuous coordinates (*i.e.* a neighbor results from a bilinear interpolation). In this study, continuous coordinates have been used to obtain a geometrically symmetric filter. The LBP filter has the property to be grayscale invariant assuming the illumination on the image is varying globally. On the other hand, it is very sensitive to noise because it captures local relationships only.

$$LBP_{P,R} = \sum_{p=0}^{P-1} s(g_p - g_c)2^p, s(x) = \begin{cases} 1, x \geq 0 \\ 0, x < 0 \end{cases} \quad (1)$$

Since the binary code of a LBP filter is retrieved with a controlled number of neighbors P , the resulting image is often represented with its histogram. For the original LBP filter, this histogram contains 2^P bins for a given radius R . Because the size of the LBP histogram increases exponentially with P , mapping methods have been proposed to reduce the number of bins contained in a LBP histogram while adding interesting properties to the filter. Reducing the number of bins intrinsically reduce the LBP footprint in memory but may reduce its discriminative power [43] if the mapping does not suit well with the data. The most popular mappings are the uniformity, $2u^2$, the rotational invariance, ri , and the combination of both, riu^2 [41]. The u^2 mapping decreases the number of bins down to $P \times (P - 1) + 3$. The riu^2 mapping decreases this number to $P + 2$. Thanks to its computational efficiency, the LBP filter has inspired many filters sharing most of its properties. Some of them are presented below in a practical way. A more detailed taxonomy of these filters has been proposed by [34].

2) *Variance Local Binary Pattern* [42]: The VAR-LBP filter consists in a LBP filter combined with the local contrast information represented by the local variance filter VAR. Since the LBP filter is grayscale invariant, it does not integrate the contrast information. The VAR filter and the LBP filter are considered to be complementary. Once the local variance has been computed at each pixel location through a circular neighborhood, a histogram of 128 bins representing the global variance of the filtered image is retrieved and concatenated to the LBP histogram [43]. In the end, the VAR-LBP filter generates a unique histogram of $2^P + 128$ bins without any mapping.

3) *Center Symmetric Local Binary Pattern* [26]: The CSLBP filter considers the information contained in the neighbors g_p only. It makes use of the symmetry of the LBP neighborhood to compute the sign of the difference between center opposite neighbors. This operation is represented by Equation (2). The central pixel g_c is unused here. The CSLBP filter produces a $\frac{P}{2}$ bits code per central pixel, resulting in a histogram of $2^{\frac{P}{2}}$ bins.

$$CSLBP_{P,R} = \sum_{p=0}^{\frac{P}{2}-1} s(g_p - g_{p+\frac{P}{2}})2^p \quad (2)$$

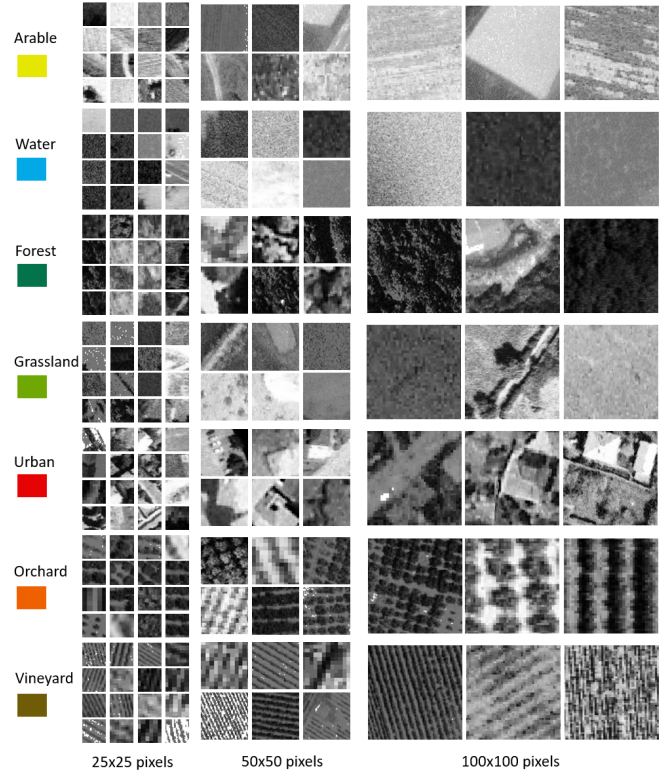


Fig. 3: Example of patches with different square sizes.

with

$$s(x) = \begin{cases} 1, x \geq \tau \\ 0, otherwise \end{cases} \quad (3)$$

and τ being a small value (*e.g.*, $\tau = 0.01$).

4) *Extended Center Symmetric Local Binary Pattern* [52]: the XCSLBP filter was presented as an enhancement of the CSLBP filter in the context of background subtraction. It has been designed to be more robust to noise than CSLBP while conserving an equivalent discriminative power. It uses intermediary metrics (see Equation (5)) in a center symmetric manner (see Equation (4)) resulting in a histogram of $2^{\frac{P}{2}}$ bins. Note that these intermediary metrics integrate the value of g_c , which is absent from the CSLBP filter.

$$XCSLBP_{P,R} = \sum_{p=0}^{\frac{P}{2}-1} s(g_1(p, c) - g_2(p, c))2^p, \quad (4)$$

where $s(x)$ is defined by Equation (3) and

$$\begin{cases} g_1(p, c) = g_p - g_{p+\frac{P}{2}} + g_c \\ g_2(p, c) = (g_p - g_c) \times (g_{p+\frac{P}{2}} - g_c) \end{cases} \quad (5)$$

5) *Three Patch Local Binary Pattern* [61]: In the TPLBP framework, a patch C is a window of w pixels \times w pixels centered on a pixel of the LBP neighborhood. The binary code of the TPLBP filter is obtained by computing the difference between two Euclidean distances. These distances are computed between the central patch C_c and two of its neighbor patches C_p and $C_{p+\alpha}$. These patches are present on the same radius R and radially spaced from an angle of value α . In

[61], α is equal to 2, resulting in a 90 degrees angle between C_p and $C_{p+\alpha}$ when the number of neighbors P is equal to 8. The TPLBP filter applied on a grayscale image generates a histogram of 2^P bins if no mapping is used. In this study, it has been decided to use $w = 1$ and $\alpha = 2$.

6) *Four Patch Local Binary Pattern [61]*: The FPLBP filter computes the difference between two Euclidean distances obtained in a center symmetric manner through the comparison of two patches radially spaced with a α degrees angle and present on two different radii R_1 and R_2 (see Equation (6)). The value of α is generally chosen as $\frac{180}{P}$. The difference between the Euclidean distances is compared to zero to produce a $\frac{P}{2}$ bits code per central pixel, resulting in a histogram of $2^{\frac{P}{2}}$ bins.

$$FPLBP_{P,R_1,R_2,w,\alpha} = \sum_{p=0}^{\frac{P}{2}-1} s(d(C_{R_1,p}, C_{R_2,p+\alpha}) - d(C_{R_1,p+\frac{P}{2}}, C_{R_2,p+\frac{P}{2}+\alpha})) \quad (6)$$

7) *Completed Local Binary Pattern [23]*: The CLBP filter combines three complementary LBP-like filters. It is defined on the same (P, R) neighborhood than the original LBP filter. The first is the classical LBP filter, renamed as CLBP_S. The other filters are CLBP_M and CLBP_C, where M stands for the magnitude and C for the center gray level. The magnitude corresponds to the absolute value of the neighbors-center difference. It is supposed to represent a complementary information to the sign which is by definition independent from the magnitude. The binary code of the center gray level is obtained by comparing each LBP center g_c with the average gray level μ of the whole image [23]. CLBP generates a concatenated histogram of $2^{P+1} + 2$ bins if no mapping is used.

8) *Local Ternary Patterns [58]*: The LTP filter is an extension of the LBP filter. It generates a ternary code instead of a binary code. The ternary values are obtained through the application of two opposite and arbitrarily chosen thresholds on the center neighbors' difference. In order to simplify its representation and to make it less computationally expensive, the LTP code can be separated into two LBP codes: one for the positive part and one for the negative part [58]. The positive part is obtained by setting all the strictly positive values to 1 and the others to 0 while the negative part is obtained setting all the strictly negative values to 1 and the others to 0. In the end, the LTP filter generates either one histogram of 3^P bins, or two histograms of 2^P bins that could be concatenated in a unique histogram of 2^{P+1} bins.

9) *Robust Local Ternary Patterns [62]*: The RLTP filter is defined as a robust to noise version of the LTP filter. For each (P, R) neighborhood including the central pixel g_c , the mean value μ_c is retrieved. The negative and positive thresholds are then defined as opposite fractions of μ_c . RLTP generates the same kind of histograms as LTP.

10) *Soft Concave-Convex Orthogonal Combination of Robust Local Ternary Patterns [62]*: The SCCOCRLTP filter is based on the RLTP filter. It proposes to increase the number of discriminative patterns while reducing their memory footprint through the concepts of orthogonal combination

[65] and concave-convex discrimination [55]. The idea behind the orthogonal combination is that a concatenation of K histograms obtained from K orthogonal filters on a (P, R) neighborhood should represent the same information than an unique histogram obtained from a complete filter but in a more compact manner (*i.e.* $K \times 2^{P/K}$ bins vs 2^P bins).

11) *Extended Local Binary Pattern [37]*: The ELBP filter is a combination of three LBP-like filters respectively named ELBP_CI, ELBP_NI and ELBP_RD. The ELBP_CI filter represents the central pixel intensity. This intensity is compared with the mean value of the whole image to get a binary code of 2 bits. The ELBP_NI filter represents the neighbor intensities in a robust to additive gaussian noise manner [37]. For each neighborhood excluding the central pixel g_c , the local mean μ of the neighbors' intensity is retrieved. The local mean μ is then compared with each neighbor to generate a binary code. The ELBP_RD filter represents the radial difference between two neighbors at the same position g_p but localized on two different radii R_1 and R_2 . The sign of the difference between each g_{p,R_1} and g_{p,R_2} neighbors is used to create the binary code. Since both the ELBP_RD filter and the ELBP_NI filter result in a histogram of 2^P bins and that the ELBP_CI filter results in a histogram of 2 bins, the ELBP filter results in a histogram of $2^{P+1} + 2$ bins without mapping.

12) *Median Robust Extended Local Binary Pattern [35]*: The MRELBP filter was presented as an update of the ELBP filter dedicated to noisy textures classification. It wisely applies median low-pass filters before the calculation of the ELBP features. The choice of a median filter has been made by Liu et al. [35] through a qualitative comparison with the gaussian and the mean filters. The MRELBP filter performs especially well on noisy texture datasets. From a practical point of view, the MRELBP filter replaces the value of a pixel with its response to a low-pass filter centered on its location. Like in [35], the size of the median filter applied to the central pixel g_c was set equal to 3. The size of the median filter for neighbors on the different radii (1,2,3) was set to (3,3,5).

B. Proposed handcrafted filters

Most filters presented above generate high dimensional feature vectors (*i.e.* 2^P bins) for only one set of parameters. Since the use of a mapping may decrease their discriminative power, the use of such filters may either provide under-accurate results with mapping or slow learning and classification without mapping. This second situation would not be suitable for online learning (*e.g.* learning from users' inputs) schemes on computers with limited computing capabilities (*e.g.* no GPU) like the ones used by the practitioners. In order to overcome these issues by finding a suitable trade-off between discriminative power and feature vector size and thus complement the current state of the art, two novel low-dimensional filters are proposed in this study.

1) *Rotated-Corner Local Binary Pattern*: The R-CRLBP filter is a novel filter that is introduced in this paper. It has been inspired by the Binary Gradient Contours (BGC) filter [19] and by the orthogonal combination [65]. It considers the sign of the successive differences between the P neighbors present

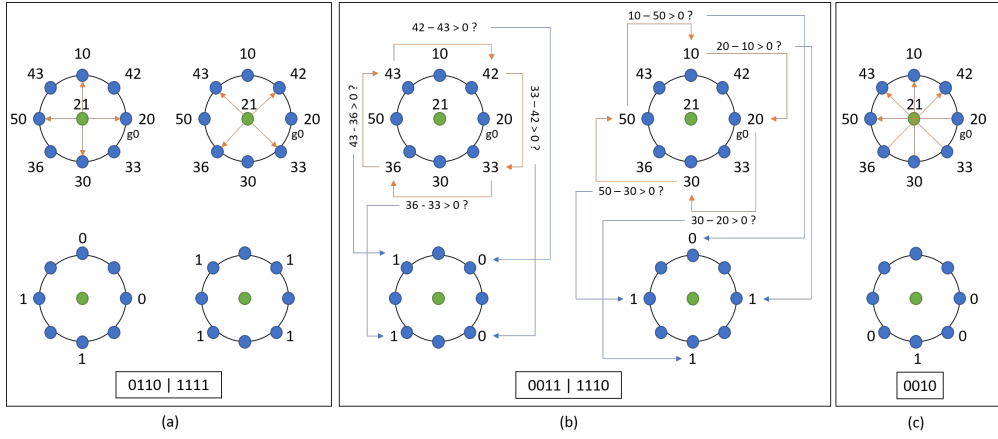


Fig. 4: Different LBP-like filters and their binary codes obtained on a same neighborhood. (a) The original LBP filter; (b) the R-CRLBP filter; (c) the CSLBP filter. The orange arrows correspond to element wise subtractions, the blue arrows show where the bit is placed, g_0 is the starting bit used to generate the binary code in a clock-wise order of counting.

on a same radius R . The successive difference between two consecutive neighbors is as follow: $(g_p - g_{p-1})$. The successive difference should be opposed to the center symmetric difference used in the CSLBP filter [26] and to the classical center neighbors difference used in the LBP filter [41] because the neighbors are neither compared symmetrically nor compared with the central pixel. From a textural pattern viewpoint, the CRLBP filter represents the *circular gradient* patterns of a neighborhood. In practice, for $P = 8$, the CRLBP filter; note the absence of the "Rotated" prescript; consists in sampling the 4 neighbors forming ± 45 degrees angles with the horizontal and vertical axis and present on a same radius R . The sign of the successive difference of the sampled neighbors is used to generate the local binary code (see Figure 4). The result of this filter is stored in a histogram of 2^4 bins. The central pixel is unused in the CRLBP filter. Multiple center rotated versions of the CRLBP filter could be used in order to consider the other neighbors in a similar manner to the orthogonal combination. Thanks to these rotations, a total of $\frac{P}{4}$ histograms could be retrieved assuming that the number of neighbors P is a multiple of 8. Each of the CRLBP rotation is computed with a unique quaternion of neighbors that is not considered in the others CRLBP rotations (see Figure 4). Note that for $P = 8$, the R-CRLBP filter is equivalent to one of the patterns used in the BGC filter. The concatenation of the histograms of the Rotated-CRLBP filters results in a unique histogram of $(\frac{P}{4}) \times 2^4$ bins. The concatenation of the CRLBP rotations will be referred as the Rotated CRLBP (R-CRLBP) filter in the rest of this paper. In Equation (7) it is assumed that the number of neighbors P is a multiple of 8.

$$R - CRLBP_{P,R} = \sum_{p=\alpha, p+=\frac{P}{4}}^{\alpha + \frac{3 \times P}{4}} s(g_p - g_{p-(\frac{P}{4})} 2^i) \quad (7)$$

with $i = \frac{p-\alpha}{\frac{P}{4}}$ and $\alpha = (0, 1, \dots, \frac{P}{4} - 1)$

2) *Light Combination of Local Binary Patterns*: LCoLBP is a combination of LBP-like filters as done in the CLBP [23], SCCORLTP [62] and ELBP [37] filters. It is formally

introduced in this paper. Its development was motivated by the will to provide a completed representation of texture patterns through an efficient low-dimensional feature vector. The LCoLBP filter consists in a concatenation of four histograms obtained with LBP-like filters, namely the FPLBP, XCSLBP, CSLBP and R-CRLBP (8). These filters have the particularity to highlight the use of the neighbors of a LBP neighborhood. Only XCSLBP uses the central pixel. The features these filters compute are based on different metrics in a sense that makes them complementary and robust to local and noisy variations. In particular, for a given neighborhood, the CSLBP and the XCSLBP histograms respectively represent the *inner gradient* patterns and the *noise robust inner gradient* patterns, while the FPLBP histogram represents the *outer gradient* patterns and the R-CRLBP histogram stands for the *circular gradient* patterns. Moreover, each of these filters output a low-dimensional histogram so that they all provide an equivalent (equal for $P = 8$) quantity of information once concatenated. Their concatenation for a given (P, R) neighborhood results in a histogram of $(\frac{P}{4}) \times 2^4 + 3 \times 2^{P/2}$ bins. In Equation (8) the function *concat* represents a 1D histogram concatenation while the *histomap* function represents the histogram retrieval operation applied on a list in a map manner.

$$histo(LCoLBP) = concat(histomap([FPLBP, R - CRLBP, XCSLBP, CSLBP])) \quad (8)$$

C. Classifiers

The classifiers used in this study are state-of-the-art supervised machine learning algorithms. They include the K-nearest neighbors (KNN) ([20], [3]), the multi-class support vector machine (SVM) ([9], [17]), the multilayer perceptrons (MLP) ([51], [50]) and the random forest (RF) classifier [11]. The parameters of these methods have been obtained via grid search as described in Section IV.

D. Learned filters and deep convolutional neural networks

Though they still lack semantic interpretation, high performing end-to-end Deep Convolutional Neural Networks (DCNNs) optimized for object recognition have been included

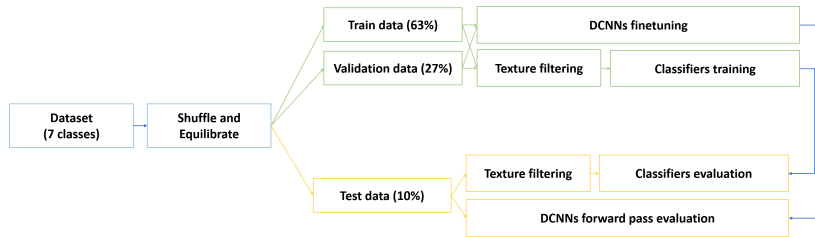


Fig. 5: The generic evaluation process performed on the HistAerial dataset for each experiment.

in this comparative study. These methods have already been successfully applied on modern satellite images [38] and they tend to outperform the handcrafted filters in classification tasks based on supervised learning. Deep filter banks have also been able to perform successfully on segmentation tasks with spatialized textured objects [14] that are structurally similar to aerial images. In contrast, theoretical [6] and experimental studies [34] have demonstrated that the DCNNs may not perform as well as expected on texture images while still providing accurate results and complementary features to the handcrafted texture descriptors [48]. Based on these contradictory results and the absence, to our knowledge, of previous studies on the DCNNs efficiency on historical aerial images, it was necessary to evaluate existing methods on the HistAerial dataset. The networks presented in this section have been selected in an iterative manner based on notable previous studies and the idea that shallower architectures should be able to at least reproduce the performances of the texture filters presented in sections III-A and III-B.

In practice, we first assessed the efficiency of LeNet [33]. We then extended our comparisons with deeper straightforward models, namely AlexNet [30] and VGG-16 [53], as well as with ResNet-18 [25] and SqueezeNet [28]. Because of its residual connections, ResNet-18 was supposed to better preserve the shallower features through the network, and thus the texture. On the other hand, SqueezeNet was presented as a network able to obtain similar results with AlexNet while being computationally less demanding at inference. When used as feature extractors, these networks generate vectors of 500 (LeNet), 4096 (AlexNet, VGG-16), 512 (ResNet-18) and 86528 (SqueezeNet)⁸ features.

IV. RESULTS AND DISCUSSION

The results for the top-1 classification (*i.e.* percentage of correct prediction considering only the prediction with the highest probability) problem were computed for the two subsets of the HistAerial dataset using the features extraction and classification algorithms presented in Section III. They are discussed below.

A. Experimental setup

1) *Handcrafted features setup*: The handcrafted filters presented in Sections III-A and III-B have been implemented in

C++ with the version 3.4 of the OpenCV library [10]. They have been applied in a same convolution manner thanks to a border warp padding strategy. In this study, the filters based on the LBP have been implemented considering a continuous neighborhood. Guided by the considerations on the computational complexity presented in [36], the values of the radius R and the number of neighbors P (see section III-A) were set to (1, 2, 3) and (8) respectively. The histograms obtained with the three (R, P) combinations were concatenated to produce a 1D vector histogram. The riu^2 mapping was applied for the filters and sub-filters producing a histogram with 2^P bins or higher to make them more practicable from a computational point of view. The riu^2 mapping was not applied on the final histogram resulting from a combination of LBP-like filters if it was already applied on the sub-filters used in this combination. As a reminder, the use of the riu^2 mapping makes sense for the HistAerial dataset since the aerial images were acquired over different years in uncontrollable conditions. As a drawback, it may result in less numerous discriminative features. No mapping was applied with the other filters. Note that the original LBP filter was evaluated with and without the riu^2 mapping in order to eventually infer a rule of thumb about its use with the historical aerial images. No preprocessing was applied on the images before the extraction of the features. The feature vectors (*i.e.* histograms) were normalized before the classification step.

2) *Classifiers and evaluation scheme*: The classifiers presented in Section III-C have been trained for each filter on the size-balanced subset of the HistAerial dataset (see Table II). Only the best processing pipelines (*i.e.* filter then classifier application) were applied on the class-balanced subset (see Table II). For both subsets, the training data were separated in train, validation and test splits as exposed on Figure 5 and Figure 6. The key hyperparameters of the classifiers were automatically obtained using a random search and then grid search strategy. The training and testing stages were both done using the Scikit-Learn library [46] (version v.0.19.1) in Python. For the KNN, K was evaluated in the range $\{1, 19\}$ with a step of 2 between two K and the Euclidean distance was used due to its generality. The SVM was trained with the RBF kernel. The other parameters for the SVM were automatically selected during the training stage looking for C among (1, 10, 100, 1000) and for γ among (0.01, 0.001, 0.0001). For the MLP classifier, the number of hidden layers for each feature vector was automatically chosen in between (1, 2, 3). The number of neurons for the first hidden layers was

⁸<https://github.com/DeepScale/SqueezeNet/issues/13>

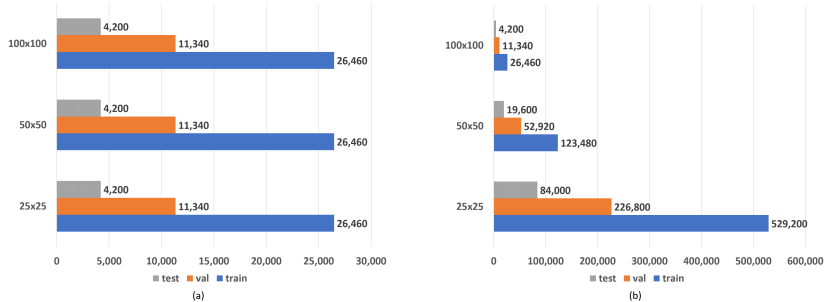


Fig. 6: Train, validation and test splits for (a) the size-balanced subset and (b) the class-balanced subset of the HistAerial dataset. Each split contains the same number of images per class. The numbers indicate the number of images.

selected as the maximum between the number of classes N_C and 0.75% of the feature vector's size S_v , as the maximum between N_C and 0.5% of S_v for the second hidden layer, and as the maximum between N_C and 0.25% of S_v for the third hidden layer. The choice to use a decreasing percentage of S_v was made as an arbitrary rule of thumb to avoid the high computational cost that a cross-validation would require on the 3 sizes of the size-balanced subset of the HistAerial dataset. The optimizer used with the MLP was the Stochastic Gradient Descent algorithm. Its best learning rate was automatically selected among (0.01,0.001,0.0001). The parameters for the Random Forest were set as 100 for the number of trees, 2 for the minimum number of samples required at a node, \sqrt{Ns} , with Ns as the number of samples, for the maximal number of samples to consider to split a node, and among (5, 10, \sqrt{Ns}) for the minimum number of samples required to split a node. The best split quality criterion was chosen between the Gini impurity and the information gain as described in [46]. All the experiments carried out with these classifiers were done on a 1.7GHz Intel i7 CPU machine with 16 Gb of available memory.

3) *DCNNs*: The end-to-end DCCNs presented in Section III-D have been trained either on the MNIST dataset (LeNet) [33] or the ImageNet dataset (AlexNet, VGG-16, SqueezeNet, ResNet-18) [44]. They have been Finetuned on a mean centered bilinearly resized versions of the HistAerial subsets (see Table II during 40 epochs. No significant improvement was shown after 40 epochs. The resizing operations were done to make the patch sizes consistent with the inputs of the DCNNs. Three channels images (*i.e.* RGB equivalent) were obtained for the ImageNet based models by stacking the same grayscale values on each color channel. While this last point is a common practice to represent grayscale images with RGB values, it may result in redundant features by approximately a factor three with DCNNs pretrained on color images because both three input channels are equal. As exposed in Section II-B the resizing operation was supposed to not modify the relative similarities between the images of the HistAerial dataset because they are square images. Both the Stochastic Gradient Descent (SGD) and the Root Mean Square Propagation (RMSPROP) algorithms were tried as optimizer algorithms for the learning part. The value of the initial learning rate has also been studied in the range of

(0.01,0.001,0.0001) to determine the best initial learning rate for each model and experiment. The learning rate decay was fixed to 0.1 and applied every 13 epochs. Both the training and testing of the DCNNs were done using the Caffe library [29] through the DIGITS application⁹ powered by three NVIDIA GPU GeForce GTX 1080 Ti. For ResNet-18, the Pytorch library [45] was used for training and evaluation due to an unavailability of pretrained weights in the Caffe library.

B. Discussion 1: global comparison

Firstly, the handcrafted filters have been compared on the Outex TC_10_000 dataset [40] to assess the correctness of their implementations. A KNN classifier using the chi square distance and $K=1$ was used. Comparable results with previous publications have been observed. The MRELBP filter combined with the riu^2 mapping, considered as a baseline on this dataset [34], achieved a score of 97.6% with $P = 8$ and $R = (1, 2, 3)$. In comparison, the LCoLBP filter achieved a score of only 51.7% with the same parameters. The score obtained with the LCoLBP filter may be explained by its non-rotational invariance definition (see Equation (8)), while the Outex TC_10_000 dataset represents rotated texture images.

Secondly, the methods have been compared on the size-balanced subset of the HistAerial dataset. In accordance with recent classification studies [53], the metric used in this study is the top-1 classification accuracy in percentage. The accuracy is defined as the percentage of well classified samples among all the test samples. In the context of multi-class classification, a top-1 accuracy score considers a sample as well classified only if its true class has been predicted as the most probable (top-1); in comparison, the commonly used top-5 accuracy score considers a sample as well classified if its true class has been predicted among the five most probable classes. Due to the high dimensional parameters space and to the multi-class classification problem considered in this study, only the best results, obtained by grid search over the classifier parameters for each filter and for each patch size, are visible on Table III, Table IV and Table V. For the patches of 25 pixels \times 25 pixels, the proposed LCoLBP filter achieved the highest score between the handcrafted filters with an accuracy of 72.9% using a MLP. The CLBP filter applied with the riu^2 mapping

⁹<https://developer.nvidia.com/digits>

TABLE III: Best results obtained on the size-balanced subset of the HistAerial dataset for the patches of 25 pixels \times 25 pixels. The missing value(s) for the the DCNNs correspond to early stops because of divergent training curves.

Handcrafted									
Filter	Parameters (P, R)	Mapping	Number of Features	Classifier - Top-1 Accuracy (%)					Rank
				KNN	SVM	RFOREST	MLP	Best	
LBP	(8,{1,2,3})	riu^2	30	65.5	62.6	67.5	64.3	67.5	9
LBP	(8,{1,2,3})	none	768	63.2	66.7	66.1	63.7	66.7	11
VARLBP	(8,{1,2,3})	riu^2	414	54.3	67.9	69.6	65.0	69.6	8
CSLBP	(8,{1,2,3})	none	48	50.4	49.3	60.8	53.1	60.8	15
XCSLBP	(8,{1,2,3})	none	48	62.5	59.1	65.9	59.0	65.9	12
TPLBP	(8,{1,2,3})	riu^2	30	61.6	56.7	62.3	59.6	62.3	14
FPLBP	(8,{1,2,3})	none	48	58.4	58.4	59.8	59.9	59.9	17
CLBP	(8,{1,2,3})	riu^2	66	69.4	69.0	72.1	68.9	72.1	4
LTP	(8,{1,2,3})	riu^2	60	66.9	65.9	69.1	69.2	69.2	7
RLTP	(8,{1,2,3})	riu^2	60	60.5	53.4	63.8	54.1	63.8	13
SCCOCRLTP	(8,{1,2,3})	none	384	52.2	54.5	54.5	50.2	54.5	20
ELBP	(8,{1,2,3})	riu^2	66	56.3	45.9	57.2	40.0	57.2	19
MRELBP	(8,{1,2,3})	riu^2	66	49.4	49.2	57.4	49.4	57.4	18
R-CRLBP	(8,{1,2,3})	none	96	63.0	65.6	65.8	66.9	66.9	10
LCoLBP	(8,{1,2,3})	none	240	68.6	71.0	71.2	72.9	72.9	3

Deep Convolutional Neural Networks									
Model	Optimizer	Epochs	Number of Features	Learning Rate - Top-1 Accuracy (%)					Rank
				0.01	0.001	0.0001	0.00001	Best	
LeNet	RMSPROP	40	500	60.0	55.3	60.2	51.7	60.2	16
AlexNet	SGD	40	4096	73.0	73.6	68.6	59.1	73.6	1
VGG-16	SGD	40	4096	—	70.3	69.9	65.8	70.3	6
SqueezeNet	RMSPROP	40	86528	—	72.6	73.1	65.2	73.1	2
ResNet-18	SGD	40	512	71.6	66.71	42.9	32.8	71.6	5

TABLE IV: Best results obtained on the size-balanced subset of the HistAerial dataset for the patches of 50 pixels \times 50 pixels. The missing value(s) for the DCNNs correspond to early stops because of divergent training curves.

Handcrafted									
Filter	Parameters (P, R)	Mapping	Number of Features	Classifier - Top-1 Accuracy (%)					Rank
				KNN	SVM	RFOREST	MLP	Best	
LBP	(8,{1,2,3})	riu^2	30	78.9	72.1	79.0	75.8	79.0	10
LBP	(8,{1,2,3})	none	768	80.5	77.9	78.9	78.5	80.5	6
VARLBP	(8,{1,2,3})	riu^2	414	67.1	77.6	80.3	78.1	80.3	8
CSLBP	(8,{1,2,3})	none	48	63.4	56.2	68.6	63.5	68.6	19
XCSLBP	(8,{1,2,3})	none	48	76.3	70.6	78.3	70.9	78.3	12
TPLBP	(8,{1,2,3})	riu^2	30	68.9	65.7	73.6	70.1	73.6	17
FPLBP	(8,{1,2,3})	none	48	72.8	70.5	74.0	71.9	74.0	16
CLBP	(8,{1,2,3})	riu^2	66	79.5	77.8	80.9	77.1	80.9	3
LTP	(8,{1,2,3})	riu^2	60	79.1	76.1	80.4	79.0	80.4	7
RLTP	(8,{1,2,3})	riu^2	60	74.4	64.2	76.6	70.8	76.6	15
SCCOCRLTP	(8,{1,2,3})	none	384	76.3	68.3	76.8	66.8	76.8	14
ELBP	(8,{1,2,3})	riu^2	66	69.1	73.7	77.9	75.0	77.9	13
MRELBP	(8,{1,2,3})	riu^2	66	65.7	61.5	71.8	65.4	71.8	18
R-CRLBP	(8,{1,2,3})	none	96	76.1	74.7	78.8	77.2	78.8	11
LCoLBP	(8,{1,2,3})	none	240	80.4	80.6	82.9	81.6	82.9	1

Deep Convolutional Neural Networks									
Model	Optimizer	Epochs	Number of Features	Learning Rate - Top-1 Accuracy (%)					Rank
				0.01	0.001	0.0001	0.00001	Best	
LeNet	RMSPROP	40	500	68.3	61.8	65.8	56.6	68.3	20
AlexNet	SGD	40	4096	82.0	82.5	78.4	68.7	82.5	2
VGG-16	SGD	40	4096	—	79.0	80.0	77.7	80.0	7
SqueezeNet	RMSPROP	40	86528	—	79.2	82.4	75.5	82.4	4
ResNet-18	SGD	40	512	82.4	74.5	60.7	37.4	82.4	3

ranked second among the handcrafted filters with an accuracy of 72.1% with a Random Forest. The LCoLBP filter resulted in a histogram of 240 bins with $P = 8$ and $R = (1, 2, 3)$. In comparison, AlexNet achieved the highest accuracy score (73.6%) with an initial learning rate of 0.001, a learning rate decay of 0.1 applied every 13 epochs, and a SGD optimizer. It generated a feature vector of 4096 bins. For the same feature vector’s size, VGG-16 achieved a best accuracy of only 70.3% with an initial learning rate of 0.0001 and the SGD optimizer. All the filter-classifier combinations obtained higher accuracy rates with wider patches. The LCoLBP filter ranked first, above the DCNNs, on the patches of 50 pixels \times 50 pixels with an accuracy of 82.9% using a Random Forest Classifier. The best DCNN obtained an accuracy of 82.5% on these data. The LCoLBP filter ranked second on the patches of 100 pixels \times 100 pixels with an accuracy of 89.3%. AlexNet ranked first on these data with a top-1 accuracy of 90.4%.

We also assessed the average execution time for features extraction for both handcrafted filters and DCNNs on a 1.7 Ghz CPU basis (*i.e.* equivalent condition with a practitioner

machine) using the optimized implementation of OpenCV 3.4 for DCNNs. We have only considered the 100 \times 100 pixels patches, since they are the most time-consuming size for handcrafted filters. DCNN are indifferent to patch size since they resize images from their original size to the expected input size. Hence, DCNNs yield the same execution time for all patch sizes. The LCoLBP filter is about 33 times faster than AlexNet. More results are visible on table V.

From an overall classification performance viewpoint, the handcrafted texture filters achieved similar results with the DCNNs while being less computationally intensive at the stages of features extraction, training and testing. In particular, the proposed LCoLBP filter achieved state-of-the-art results once combined with the Random Forest classifier. The MRELBP filter obtained surprisingly low classification rates on the HistAerial dataset compared with the Outex TC_10_000 dataset. This could be explained by the smoothing effect of the median filter applied with the MRELBP filter. The non-linearity of the median filter reduces the number of possible textural patterns that the method can generate,

TABLE V: Best results obtained on the size-balanced subset of the HistAerial dataset for the patches of 100 pixels \times 100 pixels. The missing value(s) for the the DCNNs correspond to early stops because of divergent training curves. Average (Avg.) execution time for features extraction obtained on a 1.7 Ghz CPU before classification are provided in milliseconds.

Handcrafted										
Filter	Parameters (P, R)	Mapping	Number of Features	Classifier - Top-1 Accuracy (%)					Rank	Avg.features extraction time (ms)
				KNN	SVM	RFOREST	MLP	Best		
LBP	(8, {1,2,3})	riu^2	30	87.4	81.1	87.3	83.0	87.4	9	1.047
LBP	(8, {1,2,3})	none	768	89.1	85.6	86.8	84.2	89.1	5	0.964
VARLBP	(8, {1,2,3})	riu^2	414	73.6	80.8	84.5	81.9	84.5	17	1.800
CSLBP	(8, {1,2,3})	none	48	75.7	63.2	80.3	72.8	80.3	18	0.624
XCSLBP	(8, {1,2,3})	none	48	84.4	78.2	86.0	77.5	86.0	11	0.8124
TPLBP	(8, {1,2,3})	riu^2	30	72.5	71.0	80.1	73.7	80.1	19	1.310
FPLBP	(8, {1,2,3})	none	48	84.7	79.7	85.2	81.3	85.2	15	1.023
CLBP	(8, {1,2,3})	riu^2	66	85.8	85.4	88.1	84.9	88.1	6	2.701
LTP	(8, {1,2,3})	riu^2	60	87.6	83.6	88.0	83.5	88.0	7	3.891
RLTP	(8, {1,2,3})	riu^2	60	83.6	69.3	85.3	78.1	85.3	14	2.338
SCCOCRLTP	(8, {1,2,3})	none	384	84.6	73.7	85.5	67.0	85.5	13	8.589
ELBP	(8, {1,2,3})	riu^2	66	73.9	81.8	84.8	80.5	84.8	16	3.180
MRELBP	(8, {1,2,3})	riu^2	66	74.8	82.2	85.9	79.6	85.9	12	3.528
R-CRLBP	(8, {1,2,3})	none	96	85.6	82.2	86.7	84.6	86.7	10	1.053
LCOLBP	(8, {1,2,3})	none	240	88.4	86.8	89.3	85.8	89.3	2	3.491

Deep Convolutional Neural Networks										
Model	Optimizer	Epochs	Number of Features	Learning Rate - Top-1 Accuracy (%)					Rank	Avg.features extraction time (ms)
				0.01	0.001	0.0001	0.00001	Best		
LeNet	RMSPROP	40	500	72.3	69.2	72.1	64.4	72.3	20	0.675
AlexNet	RMSPROP	40	4096	—	86.9	90.4	89.7	90.4	1	99.610
VGG-16	RMSPROP	40	4096	—	—	87.8	89.1	89.1	4	1256.500
SqueezeNet	RMSPROP	40	86528	—	86.0	89.2	84.3	89.2	3	60.772
ResNet-18	SGD	40	512	87.8	82.9	72.0	45.7	87.8	8	144.633

TABLE VI: Best results obtained on the class-balanced subset of the HistAerial dataset (*i.e.* same proportion of samples per patch size).

Best Handcrafted methods and DCNNs						
Filter	Parameters (P, R)	Mapping	Number of Features	Top-1 Accuracy (%) per patch size (pixels)		
				25 \times 25	50 \times 50	100 \times 100
LCoLBP + Random Forest	(8, {1,2,3})	none	240	75.0	84.1	89.3
AlexNet + SGD	learning rate: 0.001	*	4096	73.4	85.6	*
AlexNet + RMSPROP	learning rate: 0.0001	*	4096	*	*	90.4

which could lead to a loss in accuracy. Its negative impact is particularly visible on smaller patches, which are prone to contain fewer high frequencies than larger patches. This hypothesis is reinforced by the results obtained with the LTP filter and its robust to noise version, the RLTP filter. It is not verified by the XCSLBP filter when compared to the CSLBP filter because the XCSLBP filter does not explicitly use a low-pass filter to be more robust to noise but rather an intermediary representation of the neighborhood. Moreover, as stated in the previous paragraph, the results obtained by VGG-16, ResNet-18 and SqueezeNet on the size-balanced subset of the HistAerial dataset are lower than the results obtained by AlexNet for each patch size while other networks are deeper. These unexpected relative results would require further experiments to be investigated thoroughly, which is not the purpose of this study. However, based on the work of the authors of [59], we can make the *assumption* that DCNNs would naturally behave as low-pass filters resulting in reduced efficiency on texture data. Hence, deeper networks would generate smoother feature maps than shallower networks resulting in lower performance on texture datasets. The riu^2 mapping applied on the LBP filter did not provide any significant improvement nor significant loss on the size-balanced subset. Its use seems to be appropriated in order to reduce the computational cost of the LBP-like filters. Finally, it could be observed on Figures 7 and 8 that the LCoLBP filter and AlexNet taken in their best configurations on the *class – balanced* subset, whose accuracy results are detailed in the following section, obtained really different results on a per class basis. AlexNet optimized the representation of the Arable, Forest, Water and Urban classes, while the LCoLBP

filter provided higher accuracy rates for the Orchard and Vineyard patches. These results provide an insight about the features that have been learned by the DCNN compared with the handcrafted filter. They tend to show that these methods generate different and eventually complementary features to handle texture images. These observations agree with the results obtained by Qi et al. [48], explained in section III-D. However, it seems that both representations (*i.e.* DCNN and textures) have difficulties to differentiate Grassland and Crop classes. This could be explained by the similarity (*i.e.* low inter-class variability) of these textures and the absence of discriminative colors on the HistAerial dataset. In comparison, a Grassland is often represented with green colors and a Crop with variations of brown and yellows on the RGB images acquired during sunny periods.

C. Discussion 2: On the patch size selection

The results obtained on the size-balanced subset of the HistAerial dataset already provide information about the performance of each method on different patch sizes. However, as described in Section II-C, the size-balanced subset is made of an unbalanced percentage of data regarding the size of the patches. This subset alone does not permit to infer a conclusion about the importance of the spatial context with the HistAerial images. Another experiment was conducted on the class-balanced subset to overcome this issue. Only the methods which obtained the highest scores on the size-balanced subset have been evaluated on the class-balanced subset. Their results are presented on Table VI.

Firstly, it can be observed that similar accuracy rates are obtained on the class-balanced subset as on the size-balanced

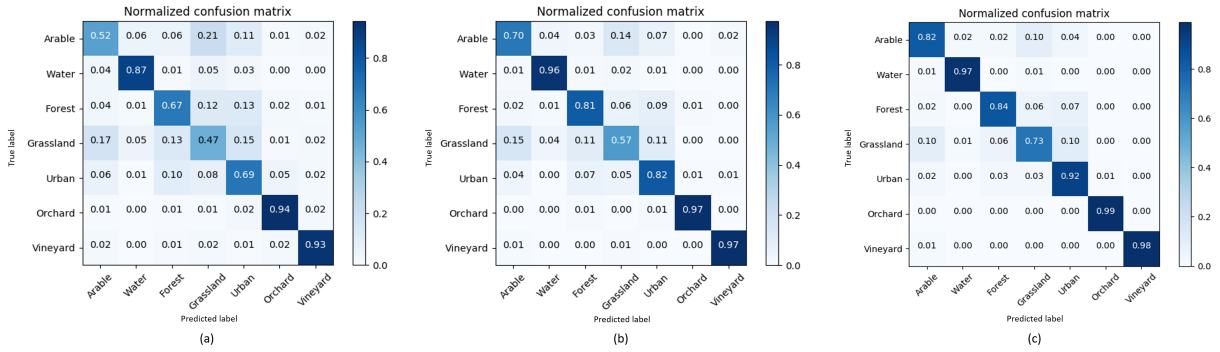


Fig. 7: Normalized confusion matrix for the best results obtained with the LCoLBP filter on the class-balanced subset. a) 25 pixels \times 25 pixels; b) 50 pixels \times 50 pixels; c) 100 pixels \times 100 pixels.

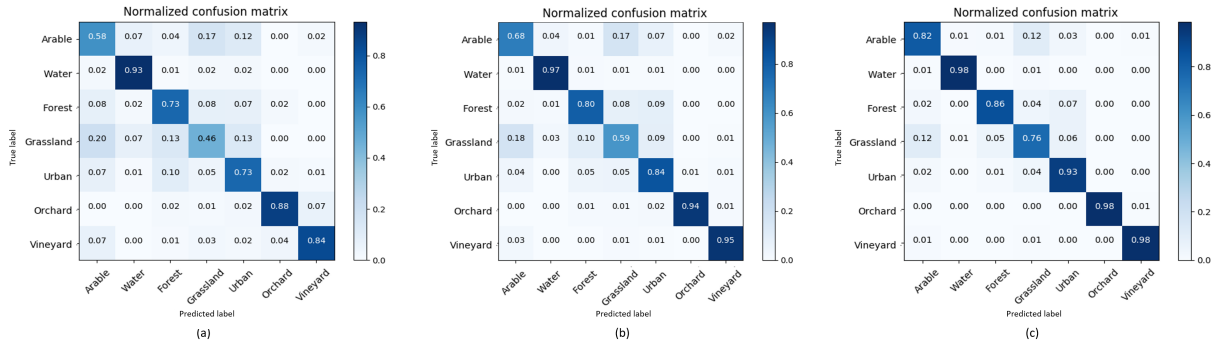


Fig. 8: Normalized confusion matrix for the best results obtained with AlexNet on the class-balanced subset. a) 25 pixels \times 25 pixels; b) 50 pixels \times 50 pixels; c) 100 pixels \times 100 pixels.

subset. The LCoLBP filter and AlexNet, respectively, achieved classification scores of 75.0% and 73.4% on patches of 25 pixels \times 25 pixels (120 thousand patches per class), and 84.1% and 85.6% on patches of 50 pixels \times 50 pixels (28 thousand images per class). These results display that a representative variability of the HistAerial dataset was already captured in the size-balanced subset.

Secondly, the LCoLBP filter performed surprisingly better than AlexNet on the patches of 25 pixels \times 25 pixels. This subset was expected to favor the deep convolutional neural network due to the large amount of data available. AlexNet did not perform as well as expected. This observation is supposed to result from the use of convolutional filters on data representing a small spatial context. This architecture obtained slightly higher scores on the other, larger, patch sizes than the LCoLBP filter with relatively short heads of 1.5% on the patches of 50 pixels \times 50 pixels and 1.1% on the patches of 100 pixels \times 100 pixels. These results are in accordance with the hypothesis presented by Basu et al. [6]: Deep convolutional neural networks seem to underperform on (non-spatialized) texture data compared with more classical real world images. Finally, the spatial context seems to provide a significant improvement (+15% in accuracy between the smallest and the biggest patches) for both handcrafted methods and DCNNs models. Since these results may have been expected for the DCNNs because they are based on spatial filters (*i.e.* convolutions), it was not trivial to guess that they would yield much higher accuracy rates with larger but less numerous

patches than with plenty of smaller patches.

V. CONCLUSIONS

In this study, an extensive comparison of state-of-the-art features extraction and classification algorithms has been carried out on a new challenging dataset made of multi-scale historical aerial images annotated by Geography experts. A novel low-dimensional filter (R-CRLBP) has been proposed. A novel Light Combination of Local Binary Patterns has been presented (LCoLBP) as an efficient low-dimensional alternative to the DCNNs for the analysis of the historical aerial images.

The proposed methods have been compared to twelve state-of-the-art handcrafted texture filters, applied with four classifiers optimized via cross-validation, and five popular deep convolutional neural networks. The results obtained for the top-1 classification task on the seven-class problem have been reported and discussed for two complementary yet randomly sampled subsets of the HistAerial dataset. The LCoLBP ranked first for each patch size between all the handcrafted texture filters. It performed equally well as the deep convolutional neural networks but with a feature vector size about 17 times shorter and an averaged features extraction speed about 33 times faster on a CPU basis. Such feature vector could allow the use of the proposed methods in interactive learning strategies (*e.g.* interaction-based learning, reinforcement learning) on non-heavily parallelized machines like the ones used by practitioners. The importance of the spatial context for the

analysis of remote sensing images through a patch-based approach has also been studied demonstrating that both the handcrafted and the DCNNs methods could achieve higher accuracy rates on the larger patches. The complementarity of the handcrafted filters and DCNNs methods has been assessed on a per-class basis.

Further work will study the combination of deep convolutional models and handcrafted texture filters for the analysis of the historical aerial images. Their eventual complementarity has been assessed in previous studies and observed in the experiments presented in this paper. The importance of the connected patches in the earth surface's classification will be studied to further assess the importance of the spatial context without requiring an extension of the proposed dataset. Finally, the use of long delay temporal series will be investigated to automate the land cover reconstruction from historical aerial images using actual multi-spectral images.

ACKNOWLEDGMENT

This work was supported by the French Environment and Energy Management Agency (ADEME) (Grant N TEZ17-42), the Rhône-Alpes Health-Environment interface platform (ENVITERA), the LABEX IMU (ANR-10-LABX-0088/ ANR-11-IDEX-0007) and the Léon Bérard Center. The authors acknowledge Etienne Durand, Geography expert, for providing the manual annotations of the historical aerial images used as the basis of this study.

REFERENCES

- [1] M A Aguilar, A Fernández, F Aguilar, F Bianconi, and A Lorca. Classification of urban areas from geoeye-1 imagery through texture features based on histograms of equivalent patterns. *European Journal of Remote Sensing*, 49:93–120, 03 2016.
- [2] A Albert, J Kaur, and M C Gonzalez. Using Convolutional Networks and Satellite Imagery to Identify Patterns in Urban Environments at a Large Scale. pages 1357–1366. ACM Press, 2017.
- [3] N S Altman. An introduction to kernel and nearest-neighbor nonparametric regression. 46:175–185, 08 1992.
- [4] O Aytekin, A Erener, I Ulusoy, and S Dugun. Unsupervised building detection in complex urban environments from multispectral satellite imagery. *Int. J. Remote Sens.*, 33(7):2152–2177, April 2012.
- [5] A L Barbieri, G F De Arruda, F A Rodrigues, O M Bruno, and L da F. Costa. An entropy-based approach to automatic image segmentation of satellite images. *Physica A: Statistical Mechanics and its Applications*, 390(3):512–518, 2011.
- [6] S Basu, M Karki, R DiBiano, S Mukhopadhyay, S Ganguly, R Nemani, and S Gayaka. A theoretical analysis of deep neural networks for texture classification. *Neural Networks*, 97, 05 2016.
- [7] R Béranger, B Blain, C Baudinet, E Faure, A Fléchon, H Boyle, V Chasles, B Charbotel, J Schuz, and B Fervers. Testicular germ cell tumours and early exposures to pesticides: The testepira pilot study. *Bulletin du cancer*, 101 3:225–35, 2014.
- [8] R Béranger, O Perol, L Bujan, F Elodie, J Blain, C Le Cornet, A Flechon, B Charbotel, T Philip, J Schz, and B Fervers. Studying the impact of early life exposures to pesticides on the risk of testicular germ cell tumors during adulthood (testis project): Study protocol. *BMC Cancer*, 14:563, 2014.
- [9] B E Boser, I M Guyon, and V N Vapnik. A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, pages 144–152. ACM, 1992.
- [10] G Bradski. The OpenCV Library. *Dr. Dobbs' Journal of Software Tools*, 2000.
- [11] L Breiman. Random forests. *Machine Learning*, 45:5–32, 2001.
- [12] M Brouwer, A Huss, M van der Mark, P Nijssen, W Mulleners, A Sas, T van Laar, G R. Snoo, H Kromhout, and R Vermeulen. Environmental exposure to pesticides and the risk of parkinson's disease in the netherlands. 107:100–110, 10 2017.
- [13] N Candade and B Dixon. Multispectral classification of landsat images: a comparison of support vector machine and neural network classifiers. In *ASPRS Annual Conference Proceedings, Denver, Colorado*, 2004.
- [14] M Cimpoi, S Maji, I Kokkinos, and A Vedaldi. Deep filter banks for texture recognition, description, and segmentation. *IJCV*, 118(1):65–94, May 2016.
- [15] P Couteron, N Barbier, and D Gautier. Textural ordination based on fourier spectral decomposition: A method to analyze and compare landscape patterns. *Landscape Ecology*, 21(4):555–567, May 2006.
- [16] N Demir, D Poli, and E Baltasvias. Detection of buildings at airport sites using images & lidar data and a combination of various methods. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38, 01 2009.
- [17] K B Duan and S S Keerthi. Which is the best multiclass svm method? an empirical study. In N C Oza, R Polikar, J Kittler, and F Roli, editors, *Multiple Classifier Systems*, pages 278–285. Springer, 2005.
- [18] A Erener and S Duzgun. A methodology for land use change detection of high resolution pan images based on texture analysis. 41:47–59, 06 2009.
- [19] A Fernández, M X Álvarez, and F Bianconi. Texture description through histograms of equivalent patterns. *J. Math. Imaging Vision*, 45(1):76–102, 2013.
- [20] E Fix and Joseph L Hodges Jr. Discriminatory analysis-nonparametric discrimination: consistency properties. 1951.
- [21] Map Projections for Europe. <http://www.ec-gis.org/sdi/publist/pdfs/annoni-et-al2003eur.pdf>.
- [22] C Gonzalo-Martin, A Garcia-Pedrero, M Lillo-Saavedra, and E Menasalvas. Deep learning for superpixel-based classification of remote sensing images. In *GEOBIA 2016*. ITC, September.
- [23] Z Guo, L Zhang, and D Zhang. A completed modeling of local binary pattern operator for texture classification. *IEEE TIP*, 19(6):1657–1663, 2010.
- [24] R M Haralick, K Shanmugam, et al. Textural features for image classification. *IEEE Transactions on Syst. Man. Cybern.*, (6):610–621, 1973.
- [25] K He, X Zhang, S Ren, and J Sun. Deep residual learning for image recognition. In *CVPR*, pages 770–778, June 2016.
- [26] M Heikkila, M Pietikainen, and C Schmid. Description of interest regions with center-symmetric local binary patterns. In *ICVGIP*, volume 4338, pages 58–69. Springer-Verlag, 2006.
- [27] X Huang, X Liu, and L Zhang. A multichannel gray level co-occurrence matrix for multi/hyperspectral image texture representation. *Remote Sensing*, 6(9):8424–8445, 2014.
- [28] F N Iandola, S Han, M W Moskewicz, K Ashraf, W J Dally, and K Keutzer. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and 1/100 model size. *ICRL*, 2017.
- [29] Y Jia, E Shelhamer, J Donahue, S Karayev, J Long, R Girshick, S Guadarrama, and T Darrell. Caffe: Convolutional architecture for fast feature embedding. *arXiv preprint arXiv:1408.5093*, 2014.
- [30] A Krizhevsky, I Sutskever, and G E Hinton. ImageNet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.
- [31] N Kussul, M Lavreniuk, A Shelestov, and B Yailymov. Along the season crop classification in ukraine based on time series of optical and sar images using ensemble of neural network classifiers. pages 7145–7148, July 2016.
- [32] N Kussul, M Lavreniuk, S Skakun, and A Shelestov. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters*, PP:1–5, 03 2017.
- [33] Y LeCun, L Bottou, Y Bengio, and P Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [34] L Liu, P Fieguth, Y Guo, X Wang, and M Pietikainen. Local binary features for texture classification: Taxonomy and experimental study. *PR*, 62:135–160, 2017.
- [35] L Liu, S Lao, P W Fieguth, Y Guo, X Wang, and M Pietikainen. Median robust extended local binary pattern for texture classification. *IEEE TIP*, 25(3):1368–1381, March 2016.
- [36] L Liu, B Yang, P Fieguth, Z Yang, and Y Wei. Brint: A binary rotation invariant and noise tolerant texture descriptor. pages 255–259, 09 2013.
- [37] L Liu, L Zhao, Y Long, G Kuang, and P Fieguth. Extended local binary patterns for texture classification. *IVC*, 30(2):86–99, 2012.
- [38] E. Maggiori. *Learning approaches for large-scale remote sensing image classification*. PhD thesis, Université Côte d'Azur, June 2017.
- [39] U Marmol. Use of Gabor filters for texture classification of airborne images and lidar data. *Archiwum Fotogrametrii, Kartografii i Teledetekcji*, 22, 2011.

- [40] T Ojala, T Mäenpää, M Pietikäinen, J Viertola, Kyllönen J, and S Huovinen. Outex - new framework for empirical evaluation of texture analysis algorithms. 2002.
- [41] T Ojala, M Pietikäinen, and T Mäenpää. *Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns*, pages 404–420. Springer, 2000.
- [42] T Ojala, M Pietikäinen, and T Mäenpää. A generalized local binary pattern operator for multiresolution gray scale and rotation invariant texture classification. In *ICAPR*, pages 399–408. Springer, 2001.
- [43] T Ojala, M Pietikäinen, and T Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE TPAMI*, 24(7):971–987, 2002.
- [44] R Olga, D Jia, S Hao, K Jonathan, S Sanjeev, M Sean, H Zhiheng, K Andrej, K Aditya, B Michael, CB Alexander, and F-F Li. ImageNet Large Scale Visual Recognition Challenge. *IJCV*, 115(3):211–252, 2015.
- [45] A Paszke, S Gross, S Chintala, G Chanan, E Yang, Z DeVito, Z Lin, A Desmaison, L Antiga, and A Lerer. Automatic differentiation in pytorch. 2017.
- [46] F Pedregosa, G Varoquaux, A Gramfort, V Michel, B Thirion, O Grisel, M Blondel, P Prettenhofer, R Weiss, V Dubourg, J Vanderplas, A Passos, D Cournapeau, M Brucher, M Perrot, and E Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [47] A Porebski, N Vandenbroucke, L Macaire, and D Hamad. A new benchmark image test suite for evaluating color texture classification schemes. *Multimedia Tools and Applications*, 70(1):543–556, May 2014.
- [48] X Qi, G Zhao, L Shen, Q Li, and M Pietikäinen. Load: local orientation adaptive descriptor for texture and material classification. *Neurocomputing*, 184:28–35, 2016.
- [49] X Qu, B Soheilian, and N Papanoditis. Landmark based localization in urban environment. *ISPRS J. Photogramm. Remote Sens.*, September 2017.
- [50] F Rosenblatt. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, 65(6):386, 1958.
- [51] D E Rumelhart, G E Hinton, Ronald J Williams, et al. Learning representations by back-propagating errors. *Cognitive modeling*, 5(3):1, 1988.
- [52] C Silva, T Bouwmans, and C Frélicot. An extended center-symmetric local binary pattern for background modeling and subtraction in videos. 2015.
- [53] K Simonyan and A Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.
- [54] I Slimene, N Chehata, J S Bailly, I Farah, and P Lagacherie. Parcel-based active learning for large extent cultivated area mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pages 1–10, 09 2017.
- [55] J Sun, G Fan, L Yu, and X Wu. Concave-convex local binary features for automatic target recognition in infrared imagery. *EURASIP Journal on Image and Video Processing*, 2014(1):23, 2014.
- [56] J Susaki and R Shibasaki. Crop field extraction method based on texture analysis and automatic threshold determination. *IGARSS*, 2:773–775, 1999.
- [57] KR Swalpa, C Bhabatosh, BC Bidyut, B Soumitro, KG Dipak, and RD Shi. Local directional zigzag pattern: A rotation invariant descriptor for texture classification. *PRL*, 2018.
- [58] X Tan and B Triggs. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE TIP*, 19(6):1635–1650, 2010.
- [59] D Ulyanov, A Vedaldi, and Lempitsky V S. Deep image prior. *CoRR*, abs/1711.10925, 2017.
- [60] JD Wegner, D Tuia, M Yang, and C Mallet. Foreword to the theme issue on geospatial computer vision. *ISPRS J. Photogramm. Remote Sens.*, January 2018.
- [61] L Wolf, T Hassner, and Y Taigman. Descriptor based methods in the wild. In *Workshop on faces in real-life images*, 2008.
- [62] X Wu, J Sun, G Fan, and Z Wang. Improved local ternary patterns for automatic target recognition in infrared imagery. *Sensors*, 15(3):6399, 2015.
- [63] L Yan and DP Roy. Automated crop field extraction from multi-temporal web enabled landsat data. *Remote Sensing of Environment*, 144:42–64, 2014.
- [64] C Zhang, X Pan, H Li, A Gardiner, I Sargent, J Hare, and P M. Atkinson. A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification. *ISPRS J. Photogramm. Remote Sens.*, August 2017.

- [65] C Zhu, C E Bichot, and L Chen. Image region description using orthogonal combination of local binary patterns enhanced with color information. *PR*, 46(7):1949–1963, 2013.



Rémi Ratajczak received his Dipl. Ing. in systems engineering from Telecom Physique Strasbourg in 2016, and his M.Sc. in robotics from University of Strasbourg, also in 2016. After graduation, he joined the Léon Bérard Centre (CLB) as an image processing engineer to support the development of innovative GIS solutions. Since fall 2017, he is pursuing a PhD in computer science in University of Lyon. His work is co-financed by the CLB and the french environment agency (ADEME), in collaboration with the LIRIS lab. His current research

topics involve texture classification and visualization enhancement applied to historical aerial images.



Carlos F. Crispim-Junior received the bachelor's degree in computer science from Universidade do Vale do Itajaí in 2006 and the Doctor degree in electrical engineering from Universidade Federal de Santa Catarina in 2011. He worked as a post-doctoral fellow at INRIA Sophia Antipolis in France from 2011 to 2017, and since September 2017 he is Associate Professor in computer science at Université Lumière Lyon 2 in Lyon, France. He is co-author of 3 book chapters, 11 publications in international journals (TPAMI, PRL, Sensors, JAD, IRBM, CompBioMed, Gerontechnology), 15 in international conferences and workshops and 1 registered, free software for video tracking and behavioral recognition on laboratory animals.



Elodie Faure holds a master's degree in geographical information systems and spatial management from Jean Monnet University (France -42), 2001. She worked as a geomatician at DRASS and then at the Rhône Alpes Regional Health Agency from 2004 to 2012. Since September 2012, she has been a geomatician in charge of GIS in the Cancer and Environment Department of the Léon Bérard Centre. She is co-author of 5 publications in international journals (Environment International, Environment Health).



Pr. Béatrice Fervers (MD, PhD, Associate Professor) is oncologist by training. She heads the Cancer and Environment Department of the Léon Bérard Cancer Centre (Lyon, France) and is affiliated to the Cancer Research Center of Lyon (CRCL, UMR Inserm 1052 CNRS 5286). She coordinates interdisciplinary research to investigate environmental exposures and their impact on cancer. Her research uses geographic information systems to characterize environmental exposures at a fine spatial scale and over long periods. From 2011 to 2017, she held the

CLARA Environment, Nutrition and Cancer Chair of Excellence, supported by Merck Serono.



Pr. Laure Tougne was born in 1971. She received her Ph.D. in 1997 in computer science from the Ecole Normale Supérieure of Lyon, France. From 1997 to 2005, she was Associated Professor at the University of Lyon 2. Since 2005, she is Professor in Computer Science at the University of Lyon 2 and she is member of the LIRIS Lab (Laboratoire d'InfoRmatique en Image et Systèmes d'information). Her research topics concern image and video analysis.