

# BOOSTING FOR INTERACTIVE MAN-MADE STRUCTURE CLASSIFICATION

Nicolas Chauffert, Jonathan Israël, Bertrand Le Saux\*

Onera - The French Aerospace Lab  
F-91761 Palaiseau, France

jonathan.israel@onera.fr, bertrand.le\_saux@onera.fr

## ABSTRACT

We describe an interactive framework for man-made structure classification. Our system is able to help an image analyst to define a query that is adapted to various image and geographic contexts. It offers a GIS-like interface for visually selecting the training region samples and a fast and efficient sample description by histogram of oriented gradients and local binary patterns. To learn a discrimination rule in this feature space, our system relies on the online gradient-boost learning algorithm for which we defined a new family of loss functions. We chose non-convex loss-functions in order to be robust to mislabelling and proposed a generic way to incorporate prior information about the training data. We show it achieves better performances than other state-of-the-art machine-learning methods on various man-structure detection problems.

**Index Terms**— Remote sensing, Machine learning, Boosting, Image classification, Object detection

## 1. INTRODUCTION

More and more satellite images are being produced, at higher and higher resolutions. The paradox is that while lots of information can now be extracted, the manual annotation of an image is so time-consuming that it prevents the deployment of such indexation schemes. Automatic processing such as segmentation and classification have been an effective way to solving this contradiction for years, but today's level of detail makes it more complicated, due to the great variety of possible visual concepts. The interactive exploration of the image is a promising way to solving this problem [1, 2, 3]. The user defines iteratively what is interesting in the image, and the system searches for areas that look similar to the selected areas. This allows us to adapt to various image types (sensor, resolution, etc.) and various geographic contexts (man-made structures do not look the same depending on which place on Earth you are).

The remainder of the paper is structured as follows. We present our interactive approach for defining the samples used for training in section 2. Features are extracted from these

samples to represent the image content (cf. section 3). Section 4 details our boosting-inspired method to perform online classification. Experimental results on real data are shown in section 5 followed by discussion in section 6.

## 2. INTERACTIVE INTERFACE FOR SAMPLE AND FEATURE EXTRACTION



**Fig. 1.** Interactive selection of regions of interest and clutter zones.

Image analysts are usually expert-users of Geographic Informations Systems (GIS) and exploit the geographic context of the image. The system in [1] keeps the complete image context visible, then learns the searched concept by using only a few selected pixels. On the contrary, Content-Based-Image-Retrieval-inspired systems for search by example (like PicSOM [2] or VisiMine [4]) segment images to small patches and display a ranked list of patches that users have to tag as good or bad. Our approach tries to mix the best of both worlds. First the user draws regions of interest and non-interest over the image using our GIS-like system named ParadisSAT (cf. Fig. 1). Secondly, the system extracts small overlapping patches from these regions, thus building the training set. This training set is then used to learn the discrimination rule of the classification method, which is finally applied locally on every patch of the whole image.

New regions of both types can be added in further interactions to iteratively refine the result. The online classifica-

\*Corresponding author.

tion method we present in section 4 only modifies the classification rule according to the new provided samples, without repeating the whole training (cf. section 5, Fig. 4). Two key-points need to be considered:

**Mislabelled data:** The interactive definition of the regions is sensitive. By drawing a region with the wrong label or selecting a larger-than-necessary area, the user may introduce mislabelled samples in the training data.

**Unbalanced data:** In lots of man-made-structure search use-cases, it is easier to find negative samples than positive ones. Thus, the user introduces a bias in the training set that eventually leads to misclassifications.

### 3. FEATURE EXTRACTION FOR MAN-MADE STRUCTURES

To feed the classification method, we investigated several state-of-the-art features [5] commonly used for man-made structure classification to represent the patches: Histograms of Oriented Gradients (HOG), Local Binary Patterns (LBP), a right angle / Line Segment Detector (LSD), edge density and SIFT. To have a good speed-performance tradeoff, we selected the fastest ones to ensure that the system has a short response time (cf. Table 1), ie. a combination of HOG and LBP. All experiments of section 5 are run with these features.

### 4. ONLINE GRADIENT-BOOSTING IN SATELLITE IMAGES

Boosting is a powerful and computationally efficient machine-learning approach. It aims at building a good (*strong*) meta-classifier from a set of weak classifiers. Several variants of the initial *adaboost* algorithm have been proposed, including the online boosting used in [7] and a more generic family of boosting methods named *online gradient-boost* [13] that builds a strong classifier  $F$  by minimizing the empirical loss defined by:

$$\mathcal{L}(F) = \sum_{n=1}^N l(y_n F(x_n)) \quad (1)$$

where  $l(\cdot)$  is a loss function (for example  $\exp(-x)$ , cf. [13] for a full list) and  $\mathcal{X} = \{(x_1, y_1), \dots, (x_L, y_L), x_i \in \mathbb{R}^D, y_i \in \{+1, -1\}\}$  is the training set of feature vectors and their associated labels.

Using online gradient-boost, we are able to propose a solution for each issue identified in section 2:

**Mislabelled data:** Boosting algorithms with a convex loss function (including the standard *adaboost*) are particularly sensitive to noise [14]. Based on the comparison of various loss functions on a man-made structure classification problem (cf. section 5), we favour the non-convex *DoomII* and *Savage* (the latter for really noisy data only) functions.

**Unbalanced data:** We propose a generic modification to the gradient-boost algorithm that consists of introducing the

prior probabilities of the training sets in the loss function  $l(\cdot)$ :

$$l(x) \leftarrow \frac{l(x)}{p(y)} \quad (2)$$

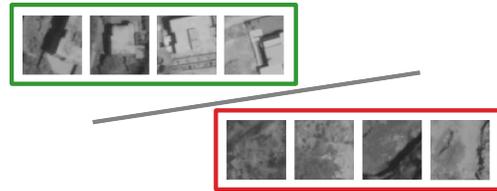
Priors are estimated using the number of positive and negative samples,  $n_+$  and  $n_-$ , by  $p(y = 1) = \frac{n_+}{n_+ + n_-}$  and  $p(y = -1) = \frac{n_-}{n_+ + n_-}$ . Consequently, the weight formulas of *online gradient-boost* in [13]-Algorithm 1 are modified in the following manner:

$$w_n = -l'(0)/p(y = y_n) \quad \text{initially} \quad (3)$$

$$w_n = -l'(y_n * F(x_n))/p(y = y_n) \quad \text{for update} \quad (4)$$

## 5. EXPERIMENTS AND RESULTS

### 5.1. Dealing with the mislabelled data issue



**Fig. 2.** Patch examples for the ground-truth dataset: man-made structures (upper row) vs. clutter samples (lower row).

We build a ground-truth dataset by extracting 50x50 patches from a 2000x2000 QuickBird image (60cm resolution) (cf. Fig. 2). It contains 615 positive samples (with houses and roads) and 1281 negative samples (woods and mountains). We compare our modified online gradient-boost (with only one iteration) with 2 state-of-the-art approaches: the standard *adaboost* (own implementation) and the support-vector machine (SVM) with a radial-basis-function kernel (libsvm implementation). Test error rates are computed using cross-validation such that reject and accept error rates are equal, by averaging the results on 10 runs. We obtain better classification rates with boosting approaches than with the SVM (cf. Table 2).

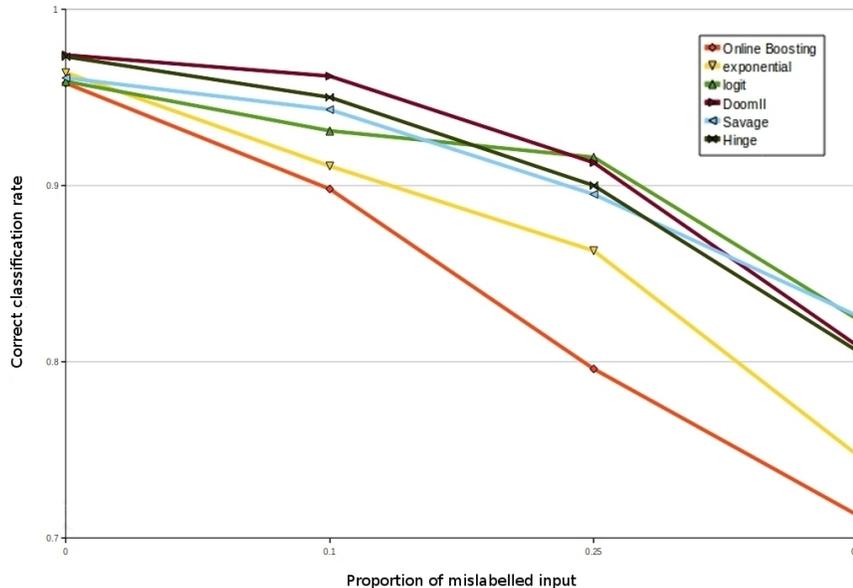
Adaboost	Online Gradient-Boost	SVM with RBF kernel
97.80%	<b>98.30%</b>	83.48%

**Table 2.** Equal Error Rates (EER) for various man-made structure classification methods in QuickBird images.

The capacity to handle mislabelled data of the various loss functions is compared by partially flipping the class of the ground-truth data. It appears on Fig. 3 that online gradient-boost with a well-chosen function is better than classic *adaboost*: with a limited noise *DoomII* has the highest performance, while with an increased mislabelling level ( $> 20\%$  of mislabelled input) *Savage* performs better.

Feature	Feature reference	Remote-sensing reference	Computation speed (s/MP)
Histograms of Oriented Gradients (HOG)	[6]	[7], [2]	1.02
Local Binary Patterns (LBP)	[8]	[7], [2]	2.39
Right-Angle Detector (LSD)	[9]	[10]	6.22
Edge Density	Canny filter	[11]	5.34
SIFT Density	[12]	[5], [10]	10.25

**Table 1.** State-of-the-art features for man-made structure classification and associated computation times.



**Fig. 3.** Influence of training-data labelling errors on performances of online gradient-boost with various loss functions.

## 5.2. Dealing with the unbalanced training data issue

To shed light on the training data balance issue, we use the ground-level-shot image dataset collected by Kumar and Hebert [15], which is highly unbalanced since there are 7 to 10 times more natural-object patches than structured ones. We compare the original loss functions and the prior-based modified loss functions. At the first iteration, prior-based DoomII catches 10 times more structured patches than the standard loss function: 887 vs. 81 patches. During the successive interactions, the user adds more natural-object regions so that the classification rate improves.

## 5.3. Real case interactive classification

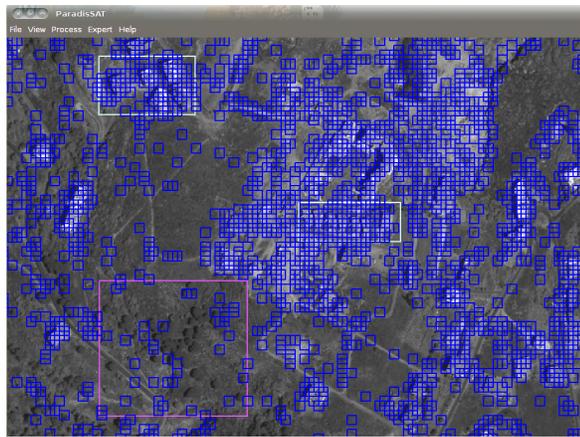
Fig. 4 shows an interactive man-made structure classification on a QuickBird image (60-cm resolution) starting from training data in Fig. 1. At interaction #2, two more areas of natural structures are added. Only two interactions yield a classification with only a few false remaining alarms, corresponding mainly to roads.

## 6. CONCLUSION

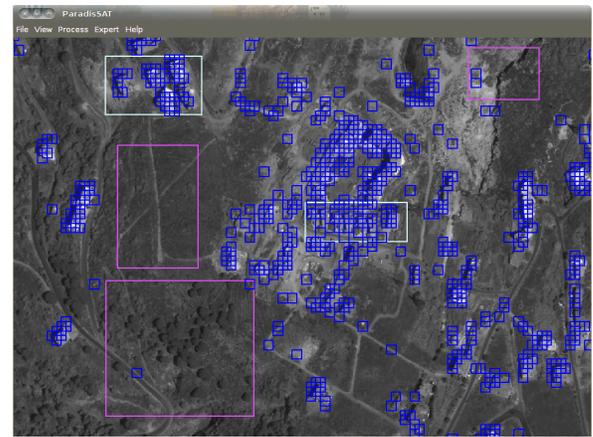
We presented an interactive framework for man-made structure detection. The key contributions are an intuitive scenario to gather training data from an image analyst, an efficient state-of-the-art feature extraction, and a redesign of the online boosting algorithm to cope with the problems raised by this context. Specifically, we assessed the suitability of online gradient-boosting with non-convex loss-functions and we proposed a generic way to incorporate prior information about the training data set into the algorithm. In the future we aim to carry out a more thorough study of the feature selection mechanism of boosting to be able to distinguish between various classes of man-made structures.

## 7. REFERENCES

- [1] M. Schröder, H. Rehrauer, K. Seidel, and M. Datcu, “Interactive learning and probabilistic retrieval in remote



(a) Iteration #1



(b) Iteration #2

**Fig. 4.** Results of an interactive man-made structure detection by online gradient-boost (first 2 rounds). Initial user inputs (turquoise blue and violet rectangles) are shown in Fig. 1. Detected locations of man-made structures are represented by 12m-width blue square areas. The detections are refined over the process to adapt to the image context.

sensing image archives,” *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2288–2298, May 2000.

- [2] M. Molinier, J. Laaksonen, and T. Häme, “Detecting man-made structures and changes in satellite images with a content-based information retrieval system built on self-organizing maps,” *IEEE Trans. on Geosci. Remote Sens.*, vol. 45, no. 4, pp. 861–874, April 2007.
- [3] Marin Ferecatu and Nozha Boujemaa, “Interactive remote-sensing image retrieval using active relevance feedback,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 4, pp. 818–826, april 2007.
- [4] K. Koperski, G. Marchisio, S. Aksoy, and S. Tusk, “Visimine: interactive mining in image databases,” in *Proc. Int. Geosci. Remote Sens. Symp.*, Toronto, Canada, 2002, vol. 3, pp. 1810–1812.
- [5] B. Sirmacek and C. Unsalan, “Urban-area and building detection using sift keypoints and graph theory,” *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1156–1167, April 2009.
- [6] N. Dalal and B. Triggs, “Histograms of oriented gradients for human detection,” in *Proc. of Computer Vision and Pattern Recognition Conf.*, Washington DC, USA, 2005, pp. 886–893.
- [7] T. Nguyen, H. Grabner, B. Gruber, and H. Bischof, “On-line boosting for car detection from aerial images,” in *IEEE Conf. on Research, Inovation and Vision for the Future*, Hanoi, Vietnam, 2007, pp. 87–95.
- [8] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 971–987, 2002.
- [9] R. Grompone von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, “Lsd: A fast line segment detector with a false detection control,” *IEEE Trans. on Pattern Analysis and Mach. Int.*, vol. 32, no. 4, pp. 722–732, April 2010.
- [10] E. Christophe and J. Inglada, “The orfeo toolbox remote sensing image processing software,” in *Proc. Int. Geosci. Remote Sens. Symp.*, Cape Town, South Africa, 2009.
- [11] X. Perrotton, M. Sturzel, and M. Roux, “Automatic object detection on aerial images using local descriptors and image synthesis,” in *Proc. of Int. Conf. on Vision Systems*, Santorini, Greece, 2008.
- [12] D. Lowe, “Distinctive image features from scale-invariant keypoints,” *Int. J. of Comp. Vis.*, vol. 60, pp. 91–110, 2004.
- [13] C. Leistner, A. Saffari, P. Roth, and H. Bischof, “On robustness of on-line boosting a competitive study,” in *Proceedings of ICCV Workshop on On-line Learning for Comp. Vis.*, Kyoto, Japan, 2009.
- [14] P. Long and R. Servedio, “Random classification noise defeats all convex potential boosters,” *Machine Learning*, vol. 78, no. 3, pp. 287–304, 2010.
- [15] S. Kumar and M. Hebert, “Man-made structure detection in natural images using a causal multiscale random field,” in *Proc. of Comp. Vis. and Pattern Rec.*, Madison, Wisconsin, 2003, pp. 119–126.