Random Subwindows and Multiple Output Decision Trees for Generic Image Annotation

Marie Dumont, Raphaël Marée, Pierre Geurts, and Louis Wehenkel

Systems and Modeling, Department of Electrical Engineering and Computer Science, University of Liège, Belgium

Abstract

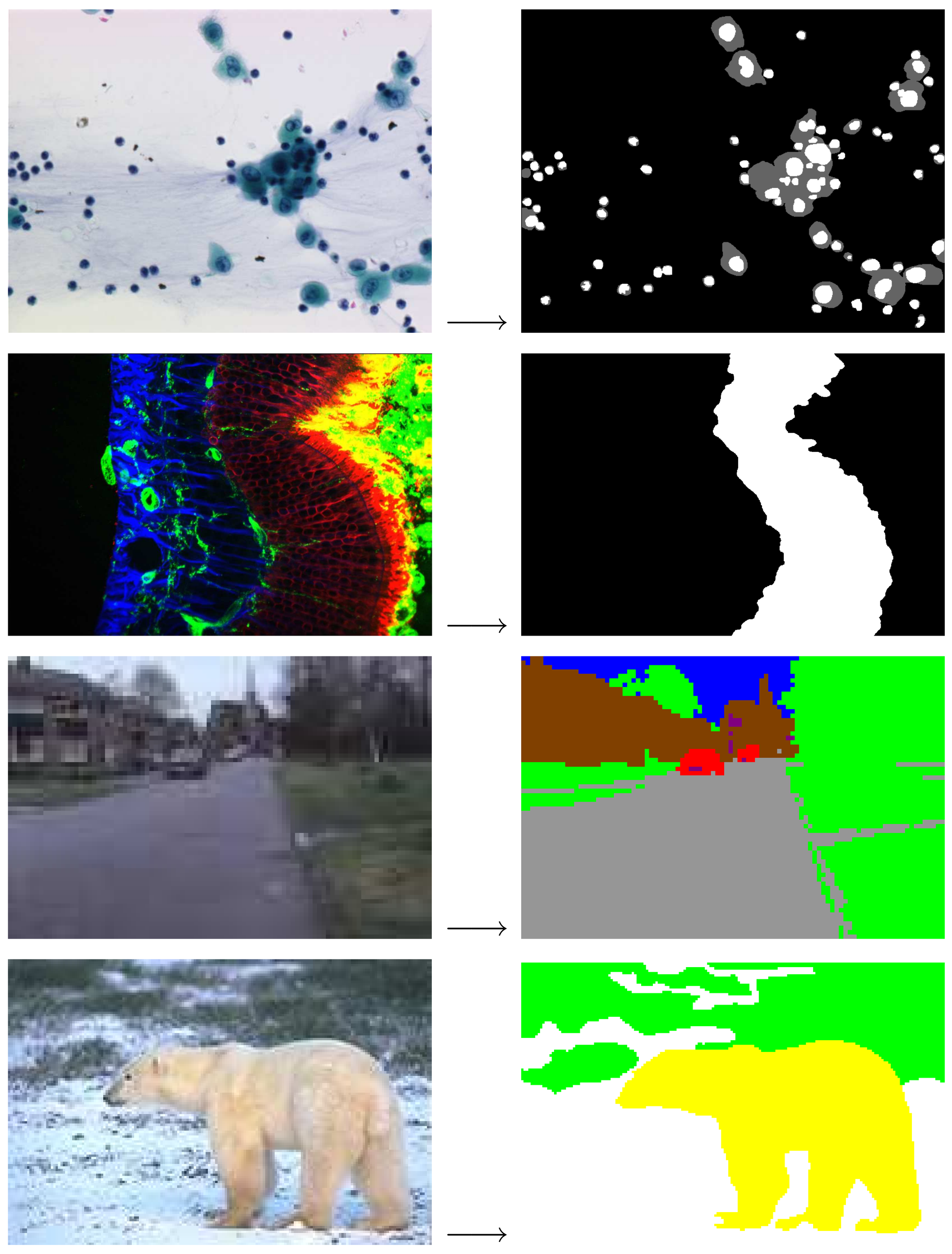
We propose a method for the generic problem of image annotation based on random subwindows extraction and ensembles of decision trees with multiple outputs. The method is evaluated on several datasets representing various types of images (microscope imaging, photographs of natural scenes, etc.).

Image Annotation

▷ Goal

Given a training set of images with pixel-wise labelling (ie. every pixel is hand-labelled with one class among a finite set of predefined classes), the goal is to build a model that will be able to predict accurately the class of every pixel of any new, unseen image. Examples on *Bronchial*, *Retina*, *Sowerby*, and *CorelA* databases:

Input image *Goal*



Background

Generic Method for Image Classification [MGPW05]

▷ Learning

- Extraction of a large number of subwindows $w_h \times w_w$ at random locations
- Subwindow described by pixels and labeled as image class
- Building a subwindow classification model using ensemble of extremely randomized decision trees [GEW06]

▷ Prediction of unseen images

- Random extraction of subwindows in test image
- Propagation of each subwindow into each tree
- Aggregation of tree votes. We assign to the image the majority class among the classes assigned to its subwindows.

Proposed Methods: an Overview

▷ Method 0: Pixel classification model

The model is built and used to predict the class of a single pixel. $\#Input = 1$, $\#Output = 1$.

▷ Method 1: Subwindow classification model

The model is built and used to predict the class of the central pixel of subwindows. $\#Inputs = w_h \times w_w, \#Output = 1$.

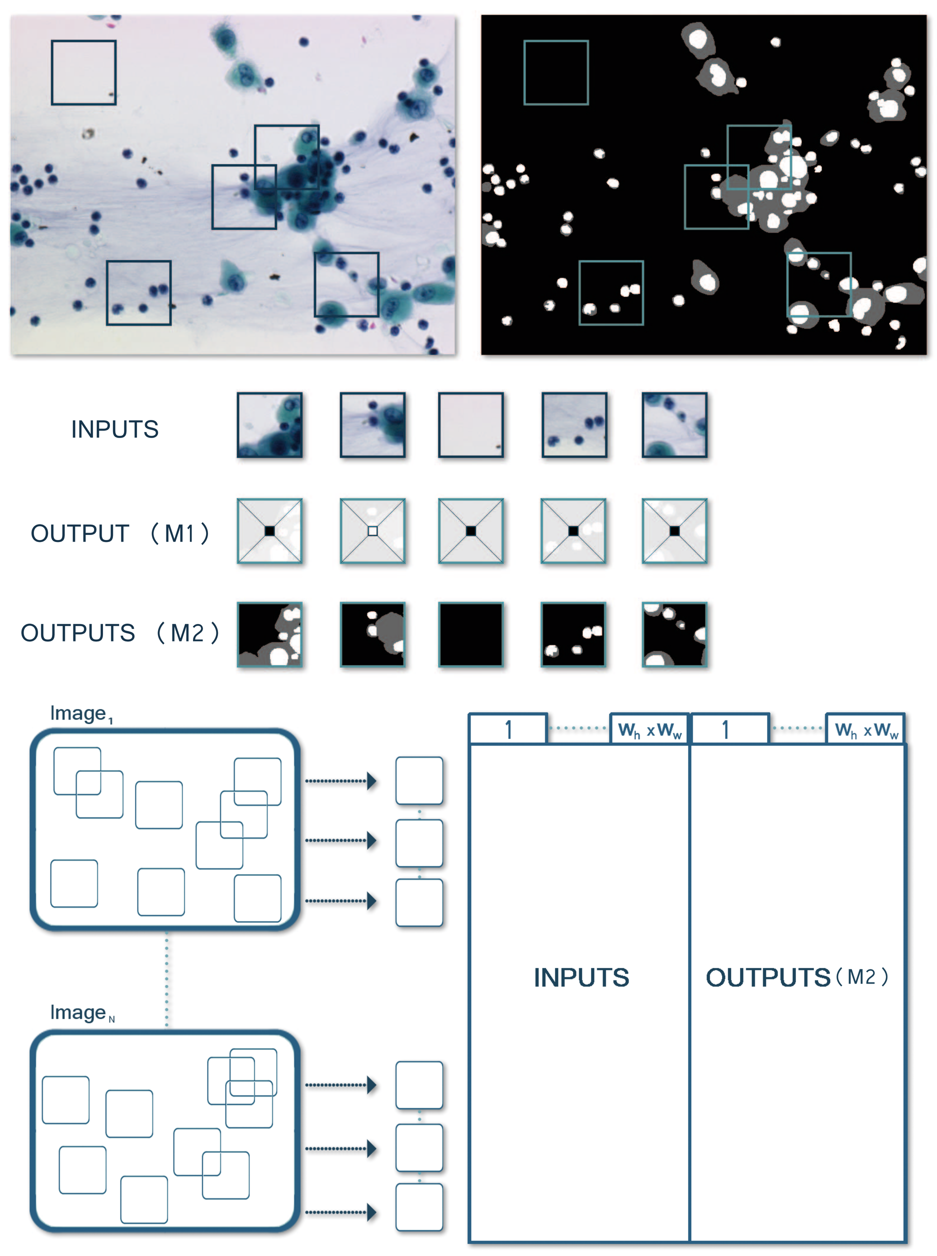
▷ Method 2: Subwindow classification model with multiple outputs

The subwindow classifier is extended so as to predict the class of every subwindow pixels.
$\#Inputs = \#Outputs = w_h \times w_w$.

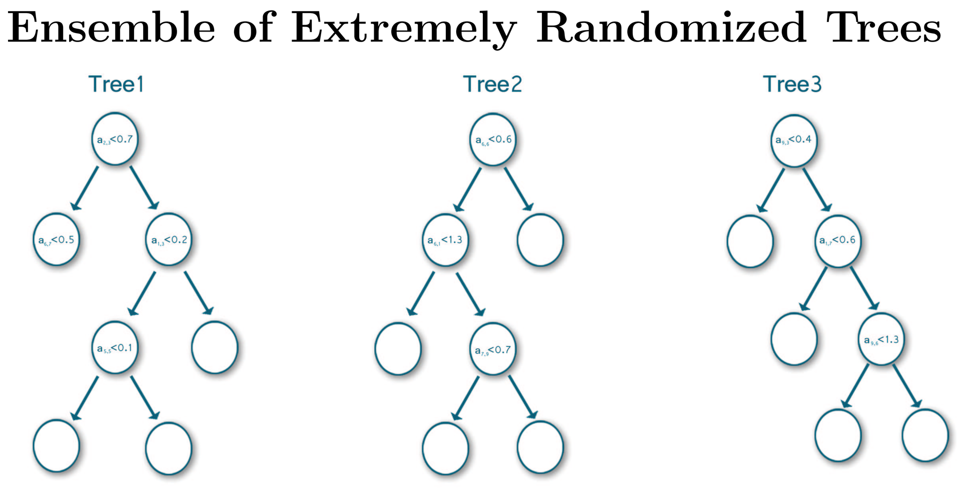
Learning and Prediction Stages

▷ Training set of Random subwindows

From the training set of N manually annotated images, fixed-size subwindows are extracted at random locations in images, described by their pixel values as inputs, and with the class color of the central pixel (Method 1) or the class colors of all pixels (Method 2) as output(s). Method 0 is a particular case with subwindows of size 1×1.



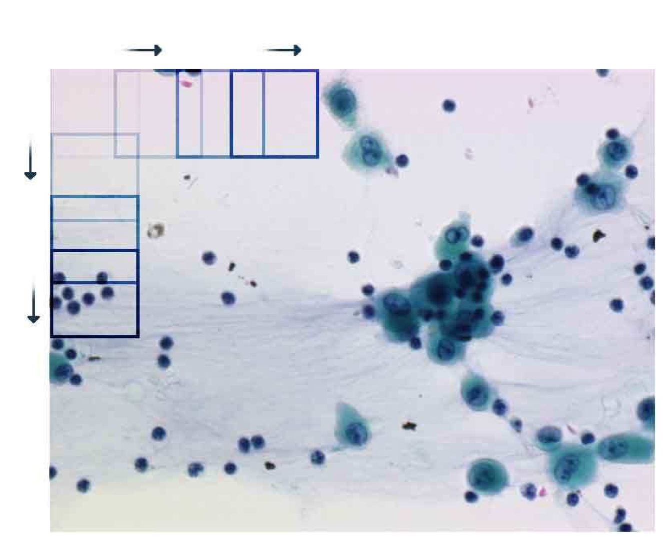
▷ Ensemble of Extremely Randomized Trees



- Top-down growing by recursive partitioning of subwindows
  - Internal *test nodes* compare a pixel component to a numerical threshold ($a_i < v_i$), *terminal nodes* output class probability estimates of the central pixel (Method 1) or information about the majority class of each output pixel (Method 2).
  - Choice of best internal tests among K random tests according to an information theoretic score computed for the central pixel class (Method 1) or average score over all output pixel classes (Method 2).
  - Fully developed (perfect fit on training set)

▷ Annotation of a new image

The model is used to classify the central pixel (Method 1) or all pixels (Method 2) of a sliding subwindow so as to annotate the full image. The sliding subwindow is moved by a step equal to w_{dist}. With Method 1, $w_{dist} = 1$. With Method 2, a class is assigned to a pixel by taking the majority class among all class predictions that were obtained for this pixel for all subwindows containing this pixel.



Results

▷ Quantitative error rates

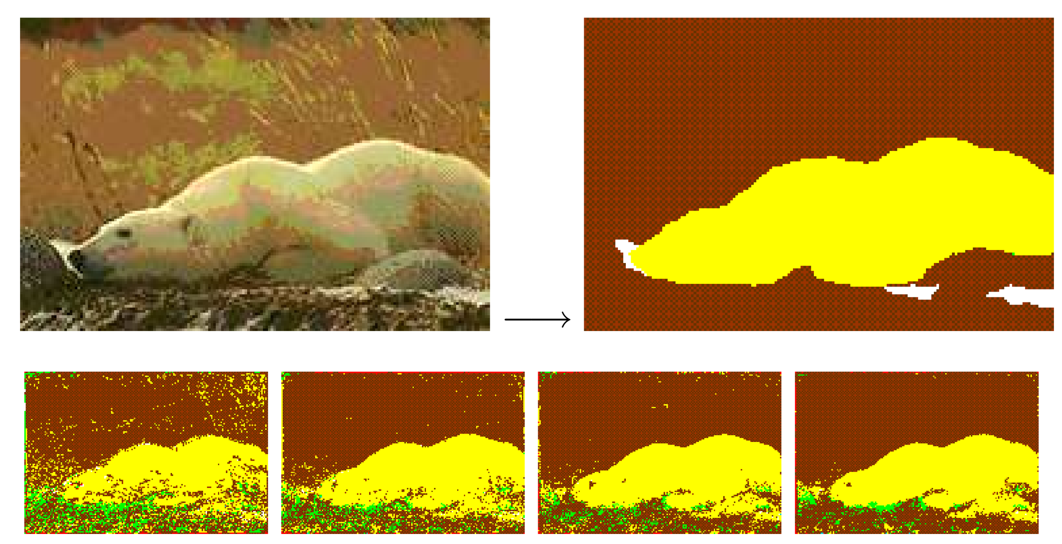
We evaluated pixel misclassification error rate (in %) compared to human annotation on five datasets, using leave-one-out protocol (first three) or independent test set protocol (last two). Best results are summarized:

Dataset	# images	# classes	M0 ER	M1 ER	M2 ER
Serous	10	2	4.07	3.61	3.28
Bronchial	8	3	4.05	3.4	3.11
Retina	50	2	13.13	5.46	5.14
Sowerby	104	8	17.07	14.86	11.02
CorelA	100	7	53.83	44.22	33.89

▷ Qualitative Observations

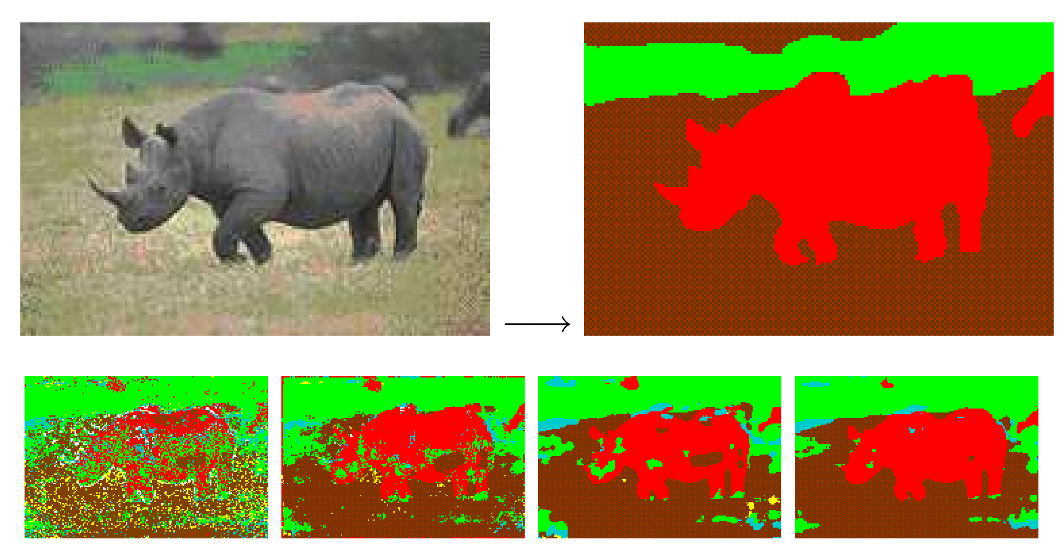
- Using more trees and more learning subwindows improve results. Illustration with increasing number of trees $T = \{1, 5, 10, 20\}$.

Input image *Manual annotation*



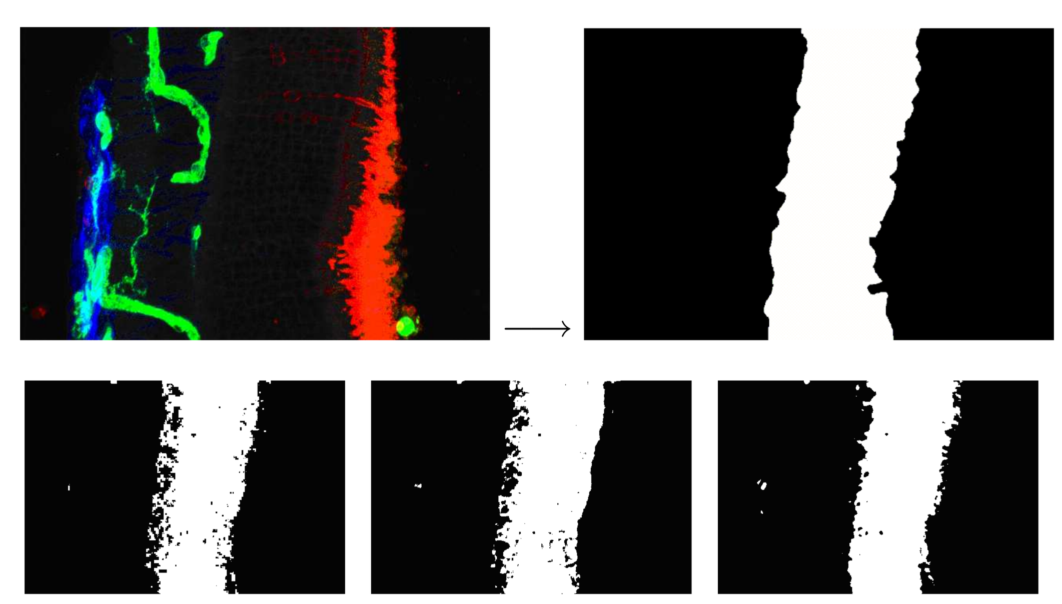
- Using multiple outputs and multiple subwindow sizes improve results. Illustration with pixel classification, subwindow central pixel output prediction, subwindow multiple output prediction, and multiple output predictions with multiple subwindow sizes.

Input image *Manual annotation*



- Using a small step for the subwindow sliding provides more precise annotation. Illustration with $w_{dist} = \{5, 2, 1\}$.

Input image *Manual annotation*



References

[GEW06] Pierre Geurts, Damien Ernst, and Louis Wehenkel. Extremely randomized trees. *Machine Learning*, 36(1):3–42, 2006.

[MGPW05] Raphaël Marée, Pierre Geurts, Justus Piater, and Louis Wehenkel. Random subwindows for robust image classification. In Cordelia Schmid, Stefano Soatto, and Carlo Tomasi, editors, *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 34–40. IEEE, June 2005.

Acknowledgments

Raphaël Marée is supported by the GIGA interdisciplinary cluster of Genoproteomics of the University of Liège with the help of the Walloon Region and the European Regional Development Fund. Pierre Geurts is a research associate of the FNRS, Belgium. Bronchial and Serous cytology images courtesy of Olivier Lezoray, Cherbourg Institute of Technology, University of Caen. Thanks to Sarah Ungar♡ for figures.