Weakly Supervised Object Detection in Artworks

CIRM: Mathematics of Imaging

Flash presentation

Nicolas Gonthier

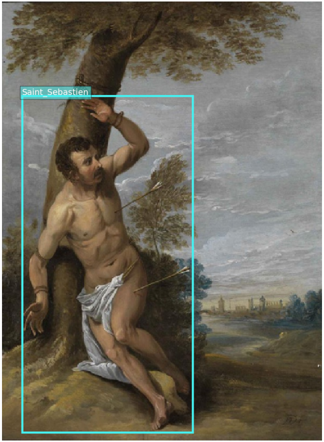
Université Paris Saclay - Télécom ParisTech - LTCI

Joint work with Y. Gousseau, S. Ladjal and O. Bonfait

7th January 2019

Motivation

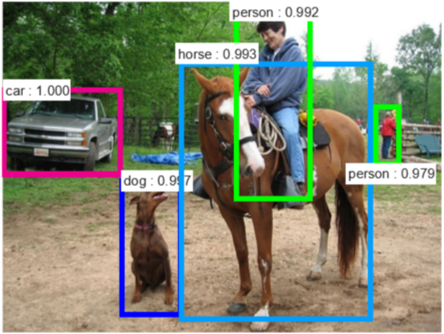
Help to search artwork databases.
We would like to **localize** the object of interest

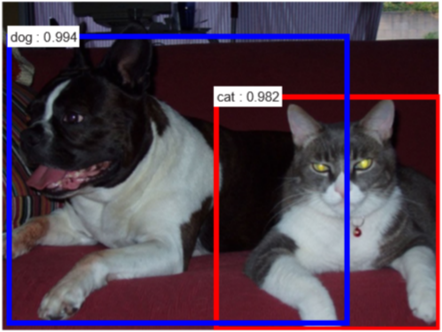


Use only **image level annotation** → **Weakly supervised** setup

Weakly supervised detection with transfer learning

Use a state-of-the-art deep network [Ren et al., 2015] **pre-trained on photography** as a feature extractor and region proposal





Source: [Ren et al., 2015]

To solve this weakly supervised problem, we use the **Multiple Instance Learning** paradigm. → Regions of an image = bag of elements

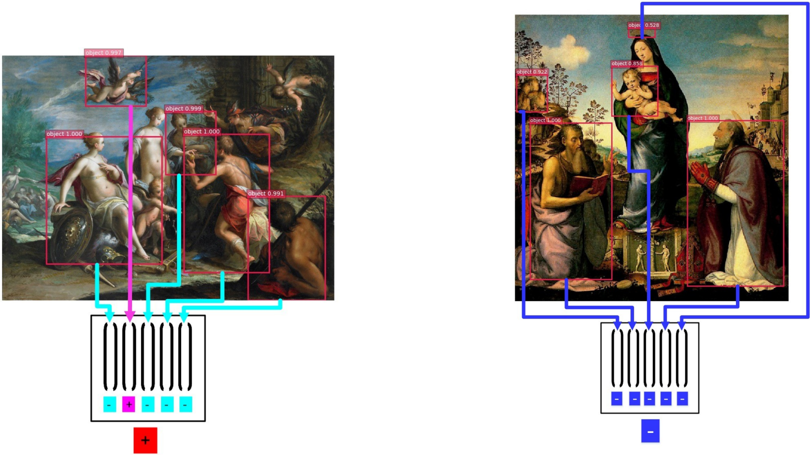


Illustration of positive and negative sets of detections (bounding boxes) for the angel category.

Model : MI-max

For each image i, we have :

$\left\{X_{i,k}\right\}_{\{1..K\}}$ features vectors

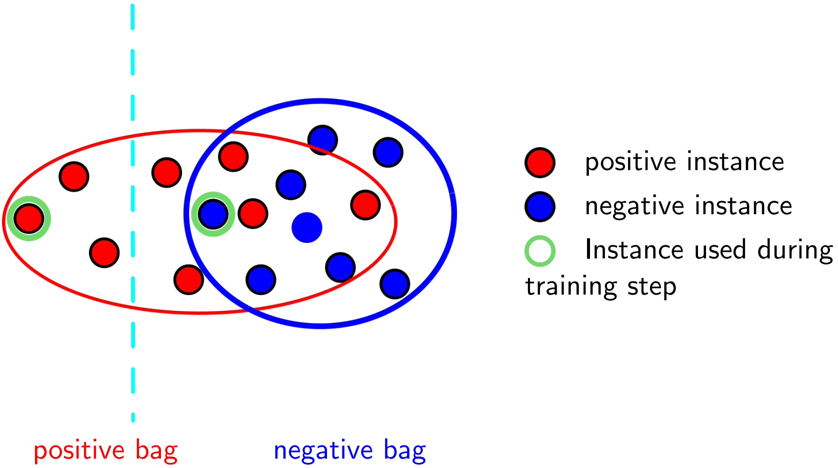
$y_i = \pm 1$ a label

We look for $w \in \mathbb{R}^M$, $b \in \mathbb{R}$ minimizing :

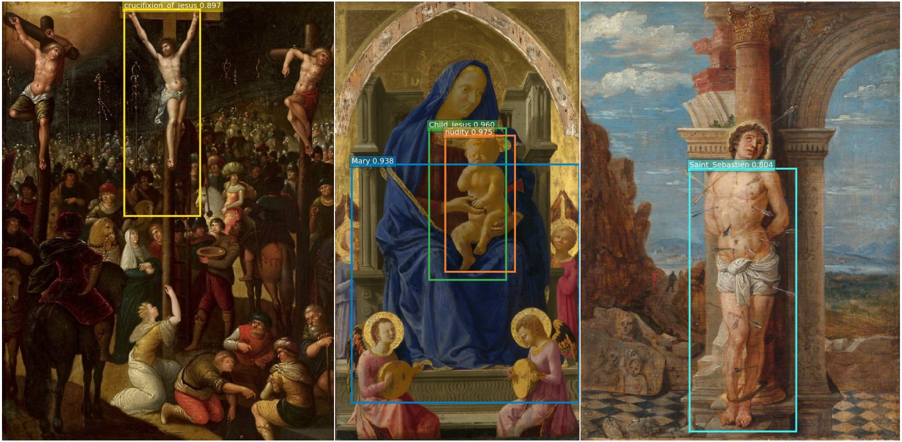
$$\mathcal{L}(w,b) = \underbrace{\sum_{i=1}^{N} \frac{-y_i}{n_{y_i}} Tanh\left\{ \max_{k \in \{1..K\}} \left(w^T X_{i,k} + b \right) \right\}}_{\text{classification loss}} \qquad \underbrace{+C * \|w\|^2}_{\text{regularisation term}} \tag{1}$$

- Experimental trick : use the objectness score produced by the detector
- Non-convex → several initialisation
- Simplified version of MI-SVM [Andrews et al., 2003] or Latent SVM [Felzenszwalb et al., 2010].

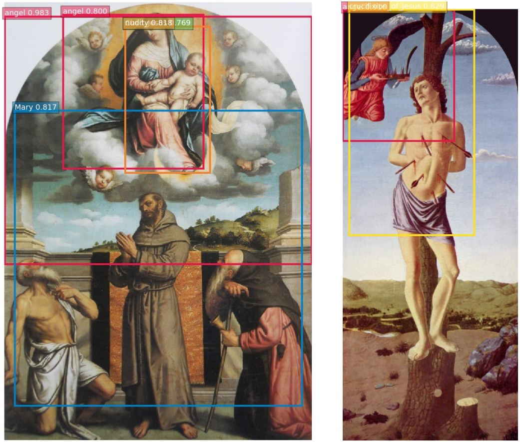
Model II : MI-max



Experiments on IconArt, successful examples



Successful examples using our MI-max-C detection scheme. We only show boxes whose scores are over 0.75.



Failure examples using our MI-max-C detection scheme.

The model is far from good.

Common Weakly Supervised problems :

- Small discriminative part of the class
- Large portion of the image

Presentation based on

Gonthier N., Gousseau Y., Ladjal S., Bonfait O.
Weakly supervised object detection in artworks,
Workshop on Computer Vision for Art Analysis, ECCV 2018
https://arxiv.org/abs/1810.02569

References I

- [Andrews et al., 2003] Andrews, S., Tsochantaridis, I., and Hofmann, T. (2003).
  Support vector machines for multiple-instance learning.
  In *Advances in Neural Information Processing Systems*, pages 577–584.
- [Felzenszwalb et al., 2010] Felzenszwalb, P. F., Girshick, R. B., McAllester, D., and Ramanan, D. (2010).
  Object detection with discriminatively trained part-based models.
  IEEE transactions on pattern analysis and machine intelligence, 32(9):1627–1645.
- [Ren et al., 2015] Ren, S., He, K., Girshick, R., and Sun, J. (2015).
  Faster r-cnn: Towards real-time object detection with region proposal networks.
  In Cortes, C., Lawrence, N. D., Lee, D. D., Sugiyama, M., and Garnett, R., editors, *Advances in Neural Information Processing Systems 28*, pages 91–99. Curran Associates, Inc.