Diverse Concept-Level Features for Multi-Object Classification

Youssef Tamaazousti, Hervé Le Borgne and Céline Hudelot

youssef.tamaazousti@cea.fr

Image Classification

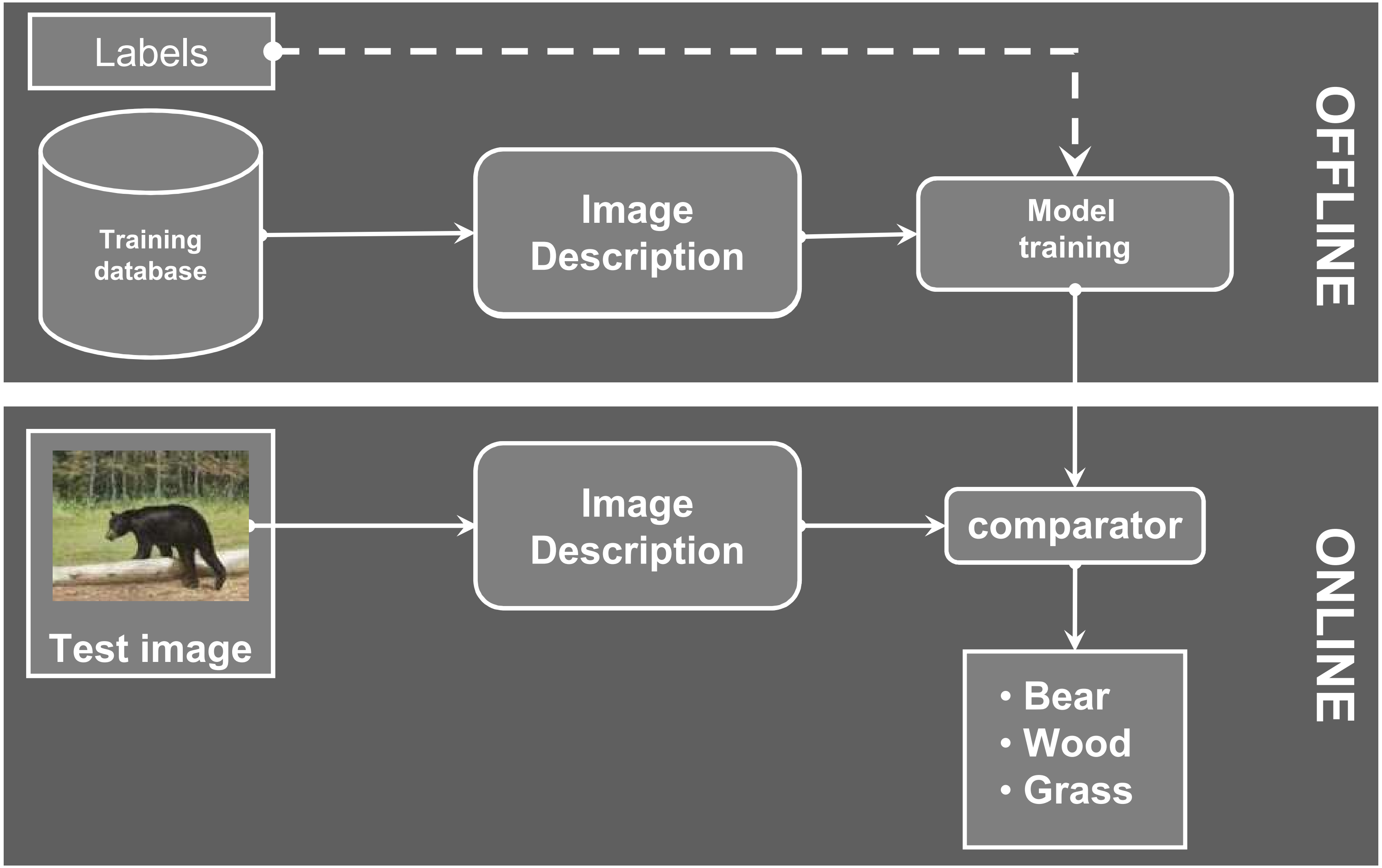


Image Classification

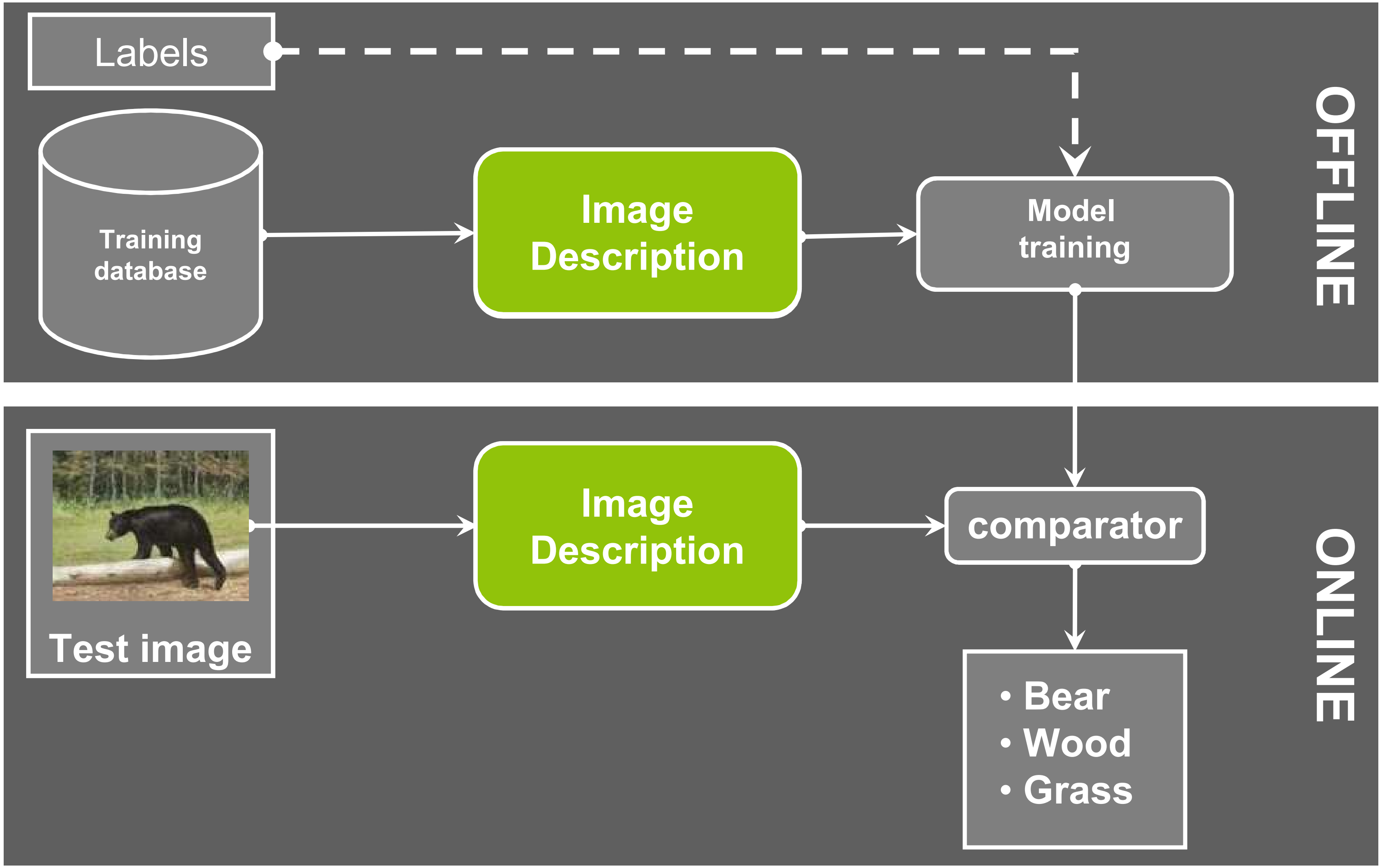


Image description

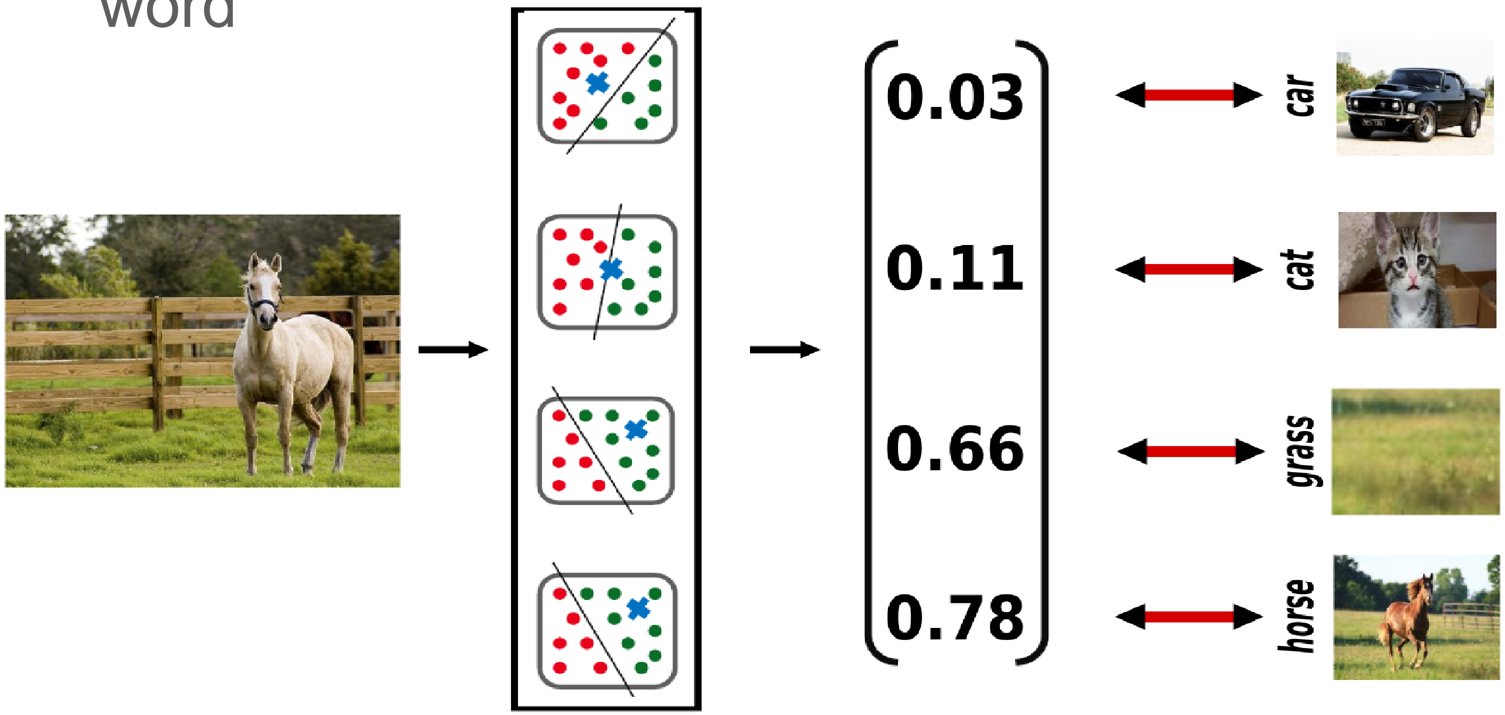
- **Low/Mid-Level Features**
  - Image described in terms of contours and shapes

- **Semantic Features**
  - Image described in terms of semantic concepts

$$\begin{pmatrix} 0.03 \\ 0.11 \\ 0.66 \\ 0.78 \end{pmatrix}$$

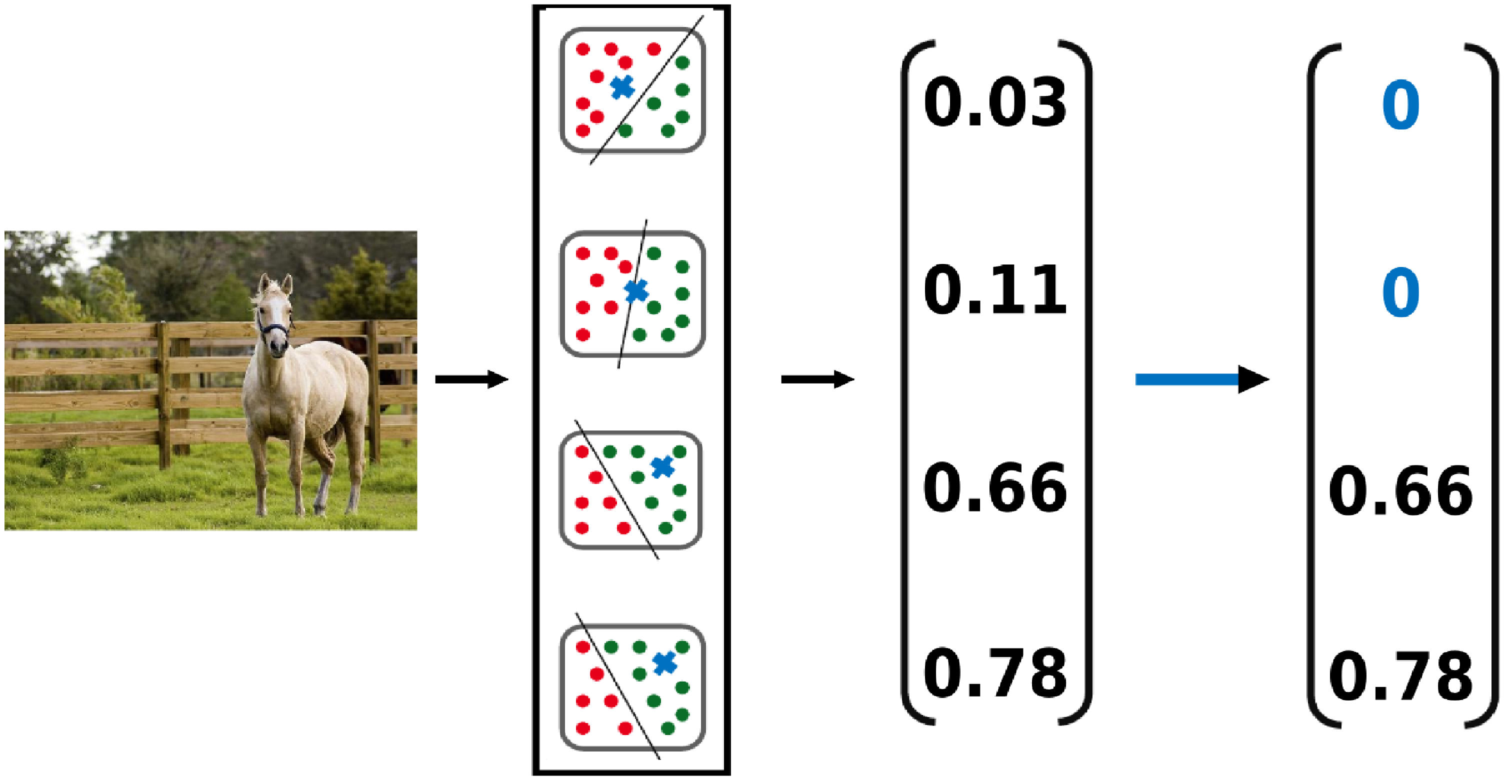
Semantic Features

- **Torresani *et al.*, 2010 – Li *et al.*, 2010**
  - Describe images in terms of outputs of concept-detectors
  - Each value is associated to a humanly-understandable word



Sparsification

- **Wang *et al.*, 2010 – Ginsca *et al.*, 2015**
  - Keep only the K highest values of the vector and set all others to zero

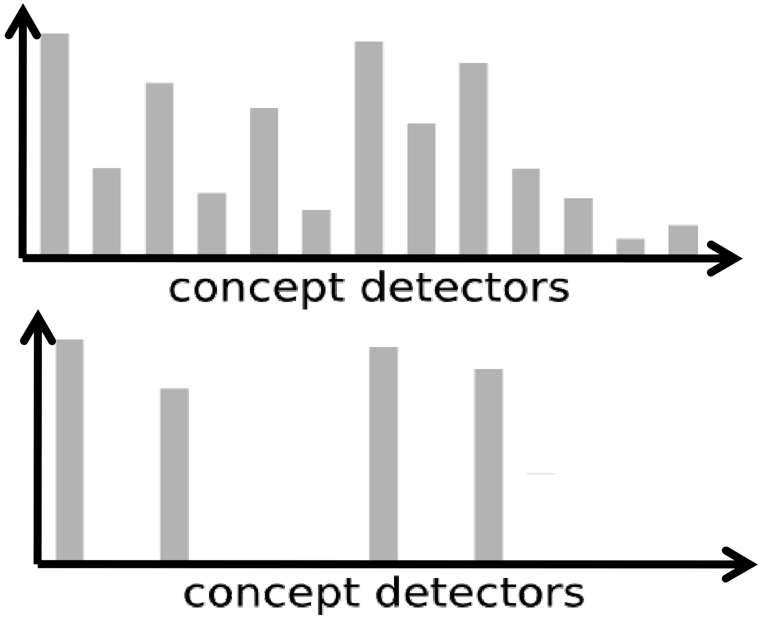


Positioning

	Sparsification: No	Sparsification: Yes
Mid-level	**Csurska et al., 2004 (Bag of Visual Words)** **Perronnin et al., 2007 (Fischer Kernels)** **Krizhevsky et al., NIPS 2012 (Fully-connected layers of CNNs)**	**Wang et al., CVPR 2010**
Semantic	**Torresani et al., CVPR 2010** **Li et al., NIPS 2010** **Bergamo et al., CVPR 2012**	**Ginsca et al., MMM 2015** **Tamaazousti et al., ICMR 2016 (Ours)**

Classificatin with Semantic Features

- **Object classification**
  - Without sparsification
    - **No missing** information but **noisy** values (not good)
  - With sparsification
    - **No missing** information (good)

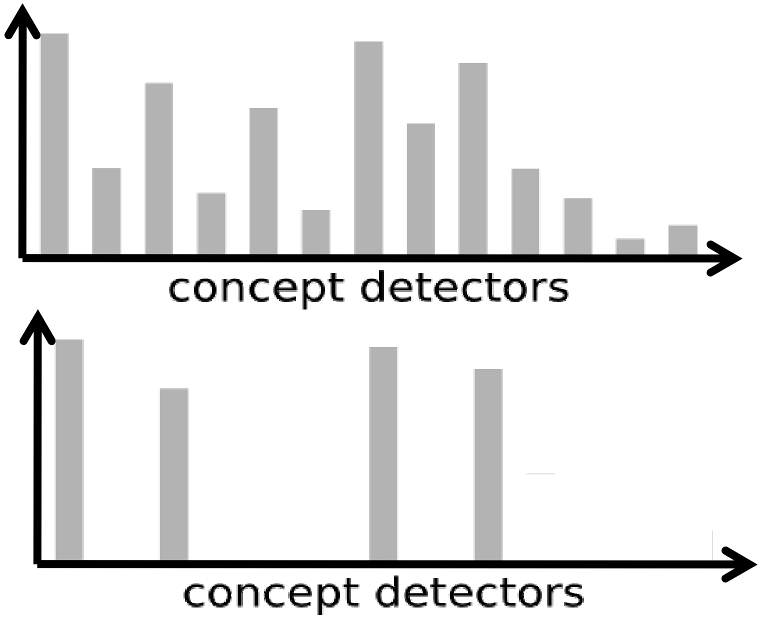


Not sparse

Sparse

Classificatin with Semantic Features

- **Multi-Object classification**
  - Without sparsification
    - **No missing** information but **noisy** values (not good)
  - With sparsification
    - **Missing** information (not good)



Not sparse

Sparse

Problem

- **Typical problematic case**
  - Image with multiple objects

- **Observation**
  - When the concept of the largest object is activated, a set of its annex concepts is also activated
- **Why are we loosing information?**
  - Naive sparsification
    - Would select one principal concept and its annex concepts
    - Other principal concepts could be set to zero

Usual formalism

- **Sparsification** [Wang *et al.*, 2010, Ginsca *et al.*, 2015]
  - Principle
    - Set to zero « some » values of the vector
  - Objective
    - Keep the **good concepts** and delete the **bad ones**
- **Usual definition**
  - Good concepts = **highest** values
  - Bad concepts = all others (**lowest** values)

Proposed formalism

- **Proposed definition**
  - Good concepts = principal concepts and their annex concepts (<u>not necessarly the highest</u> values)
  - Bad concepts = all others (<u>not necessarly the lowest</u> values)
- **Questions**
  1. **How to get the good concepts?**
  2. **What are the good concepts ?**

1. How to get the good concepts ?

- **Get the good concepts is a hard problem !**

- **Bergamo *et al.*, 2012 (Bottom-up)**
  - Get generic concepts (**good concepts**) using unsupervised clustering (hard)
  - Bottom-up: Low-level errors are propagated to upper concepts → limited performances

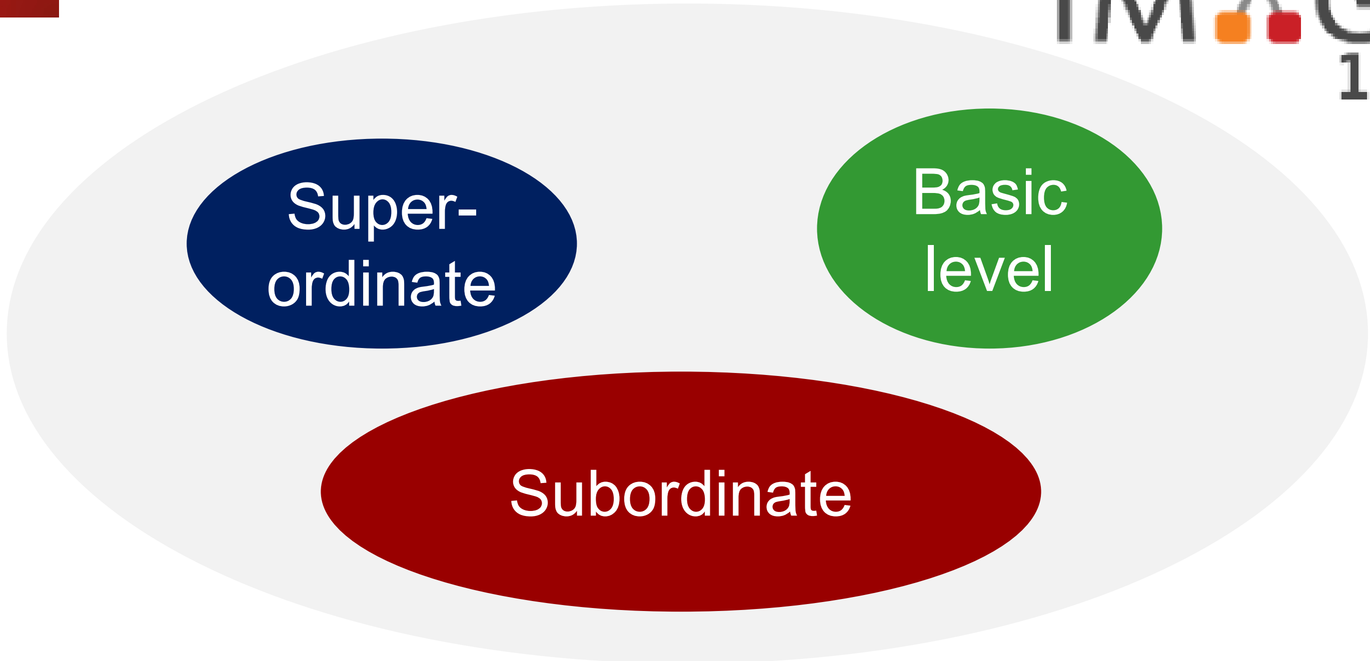
- **Our proposal (Top-Down)**
  - Get the good concepts using largely available Human Knowledge databases (hierarchies, human-categorization rules, databases, etc.)

2. What are the good concepts?

- **Inspired by Psychological studies**
- **Rosch, 1978 - Jolicoeur *et al.*, 1984**
  - Different levels of good concept in Human minds
  - The concepts mostly known and used by Humans are
    - Superordinate: **vehicle**
    - Basic-level: **car**
    - Subordinate: **ford mustang**

Observations

IMAGENET
17k concepts



	Number of concepts-detectors	Range of values of concept-detectors
Superordinate	Low	Low
Basic-level	Normal	Normal
Subordinate	High	High

Proposed approach

- **Concept-detectors**
  - **Superordinate**
    - Semantic process ⟶ High range of values
  - **Basic-level**
    - Visual process
  - **Subordinate**
    - Visual process + reduction of number of concepts
      ⟶ Low number of concepts

	Number of concepts	Range of values
Superordinate	Low	Low ⟶ High
Basic-level	Normal	Normal
Subordinate	High ⟶ Low	High

Proposed approach

- S.O.T.A semantic feature

$$\mathcal{F} = \left[* \;\cdots\; * \;\; 0 \qquad \cdots \qquad 0 \right]$$

- Our final semantic feature (D-CL)

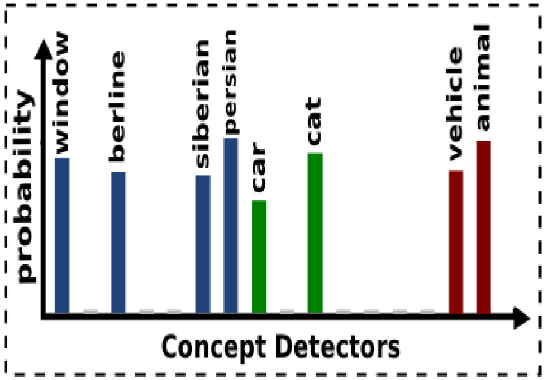
Superordinate Basic-level Subordinate

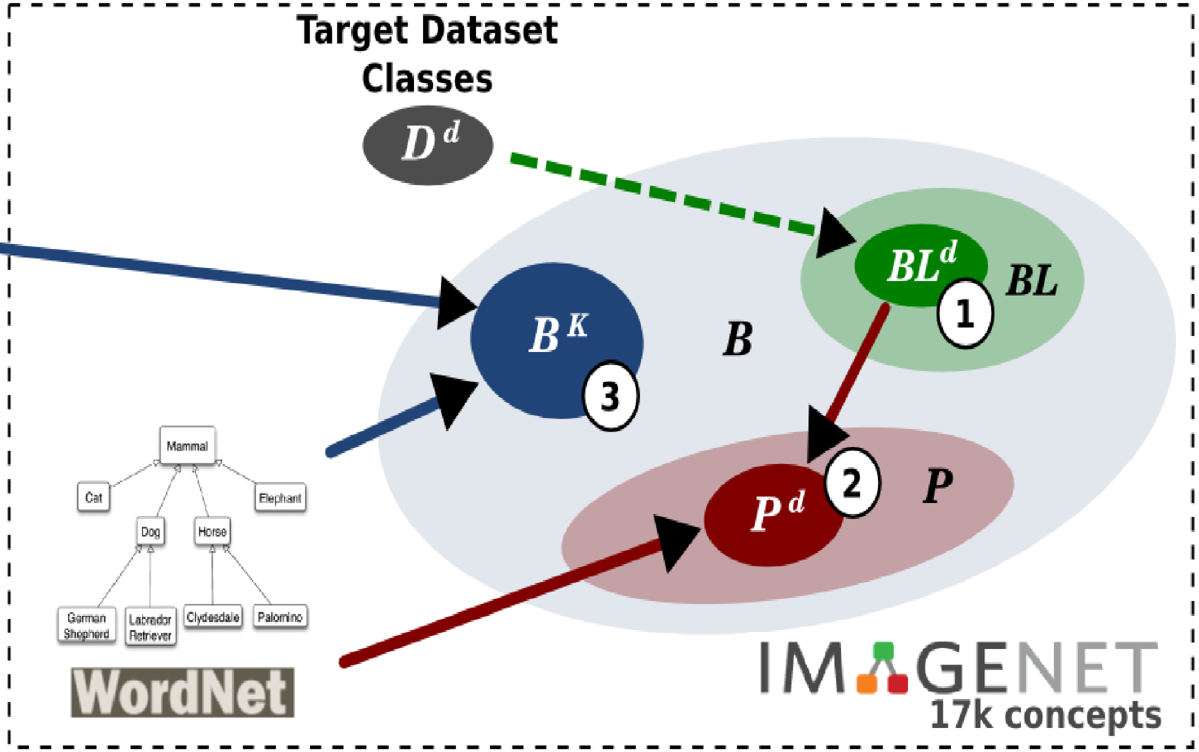
$$\mathcal{F} = \left[\underbrace{* \;\cdots\; *}_{\substack{\text{Semantic process:}\\ \text{semantic classifiers}}} \; \underbrace{* \;\cdots\; * \;\; * \cdots * \; 0 \;\cdots\; 0}_{\substack{\text{Visual process:}\\ \text{binary classifiers}}} \right]$$

In practice

- Hard to set the list of ***superordinate***, ***basic-level*** and ***subordinate*** concepts

input image





Final Semantic feature

Get the diverse levels of concepts

Experimental Protocol

	Pascal VOC 07	Pascal VOC 12	Nus-Wide Object
Benchmark	✓	✓	✓
Rate of multi-label	45%	30%	20%

- **Evaluation metric**
  - mean Average Precision (mAP)

- **Pascal VOC 07**
  - Train/val: 5k images - Test: 5k images
- **Pascal VOC 12**
  - Train/val: 10k images - Test: 10k images
- **Nus-Wide Object**
  - Train/val: 20k images - Test: 15k images

Multi-Object Classification Results

Method	Nus-Wide Object (20%)	Pascal VOC 2007 (45%)	Pascal VOC 2012 (30%)
Li *et al.*, 2010	n.a	45.2	n.a
Torresani *et al.*, 2010	n.a	43.8	n.a
Torresani *et al.* (reimpl.)	70.3	82.4	81.7
Bergamo *et al.*, 2011	n.a	43.7	n.a
Bergamo *et al.*, 2012	36.5	53.2	49.3
Simonyan *et al.*, 2015	67.3	77.4	77.2
Ginsca *et al.*, 2015	74.7	82.8	81.7
D-CL (ours)	**76.0**	**85.1**	**83.0**

Without sparsification

Naive sparsification

Conclusions

- **Novelty:**
  - New semantic image-representation
  - New formalism of sparsification
  - New sparsification process based on Human-cognition

- **Results:**
  - Multi-object classification
    - 3 publicly available benchmarks
    - +2 points of mAP compared to the best state-of-the-art semantic features

Thank you
(questions ?)

Commissariat à l'énergie atomique et aux énergies alternatives
Institut List | CEA SACLAY NANO-INNOV | BAT. 861 – PC142
91191 Gif-sur-Yvette Cedex - FRANCE
www-list.cea.fr

Établissement public à caractère industriel et commercial | RCS Paris B 775 685 019