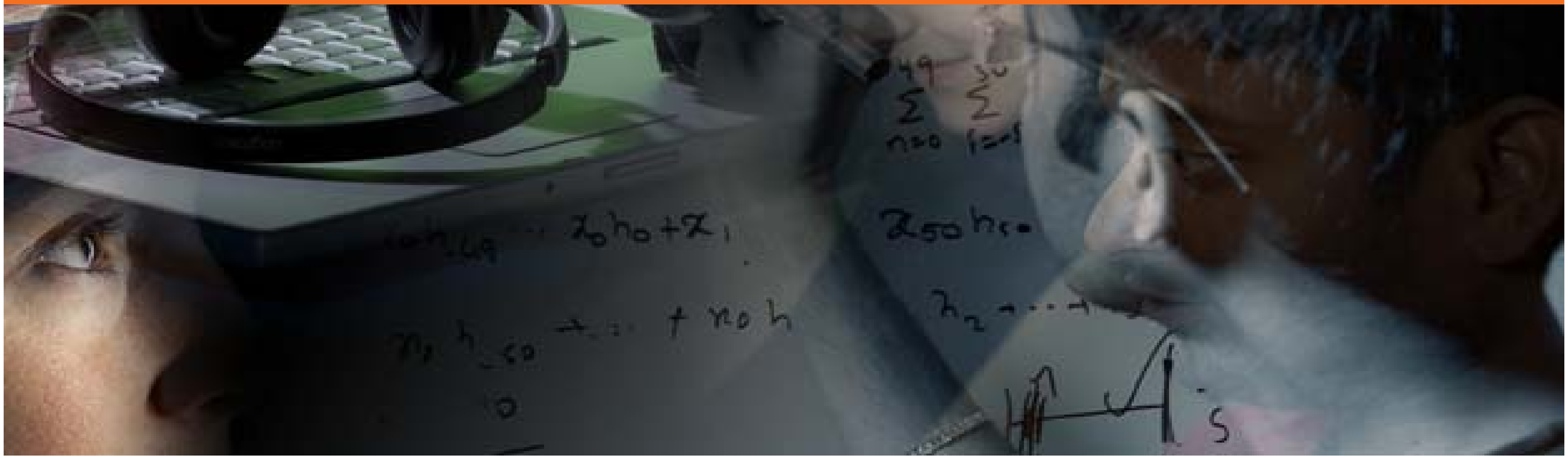
Learning and Sharing of Semantic Spatial Concepts

Barbara Caputo

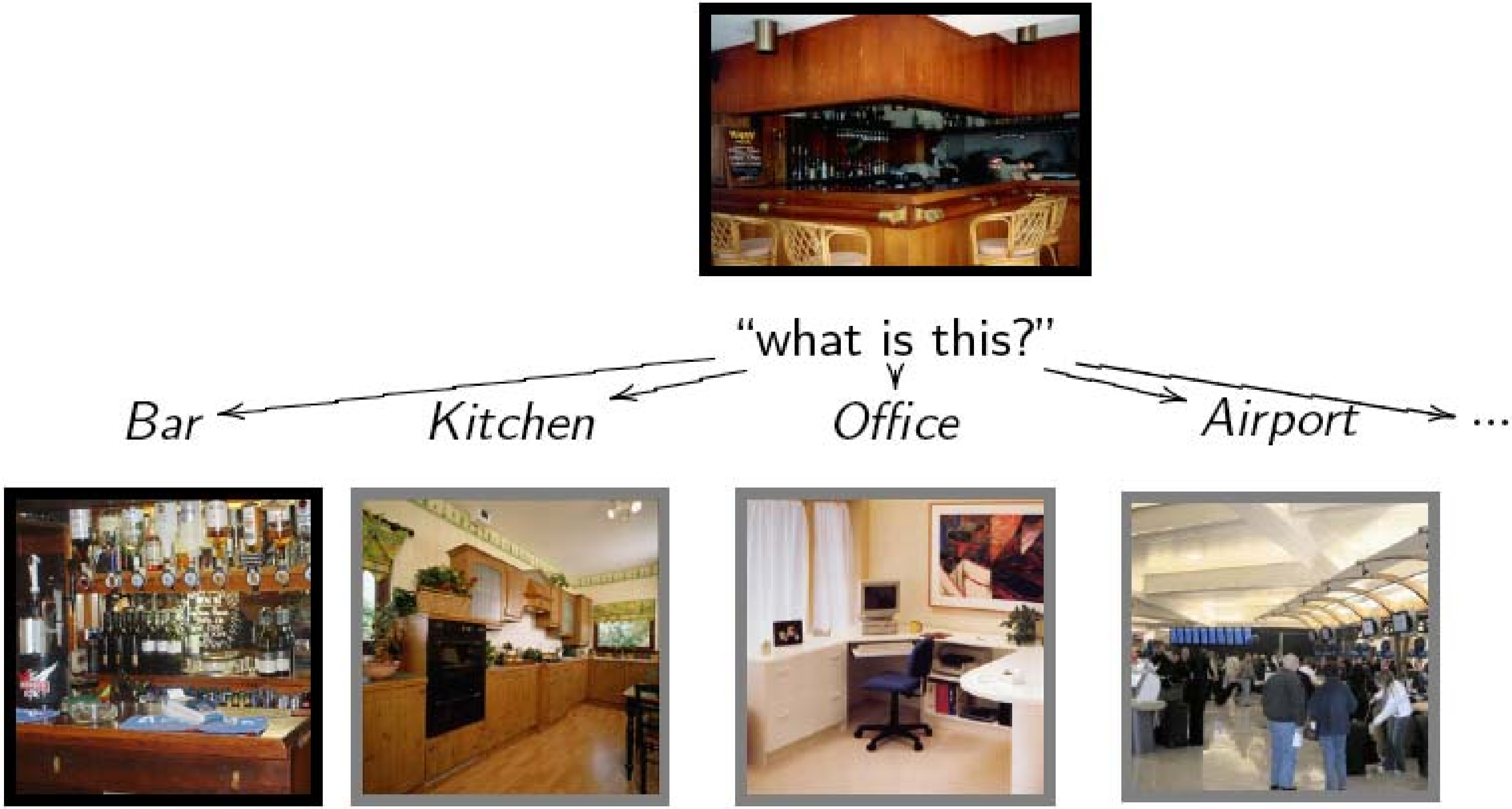
International Advisory Board Meeting, Idiap, September 2, 2011

Research Activities



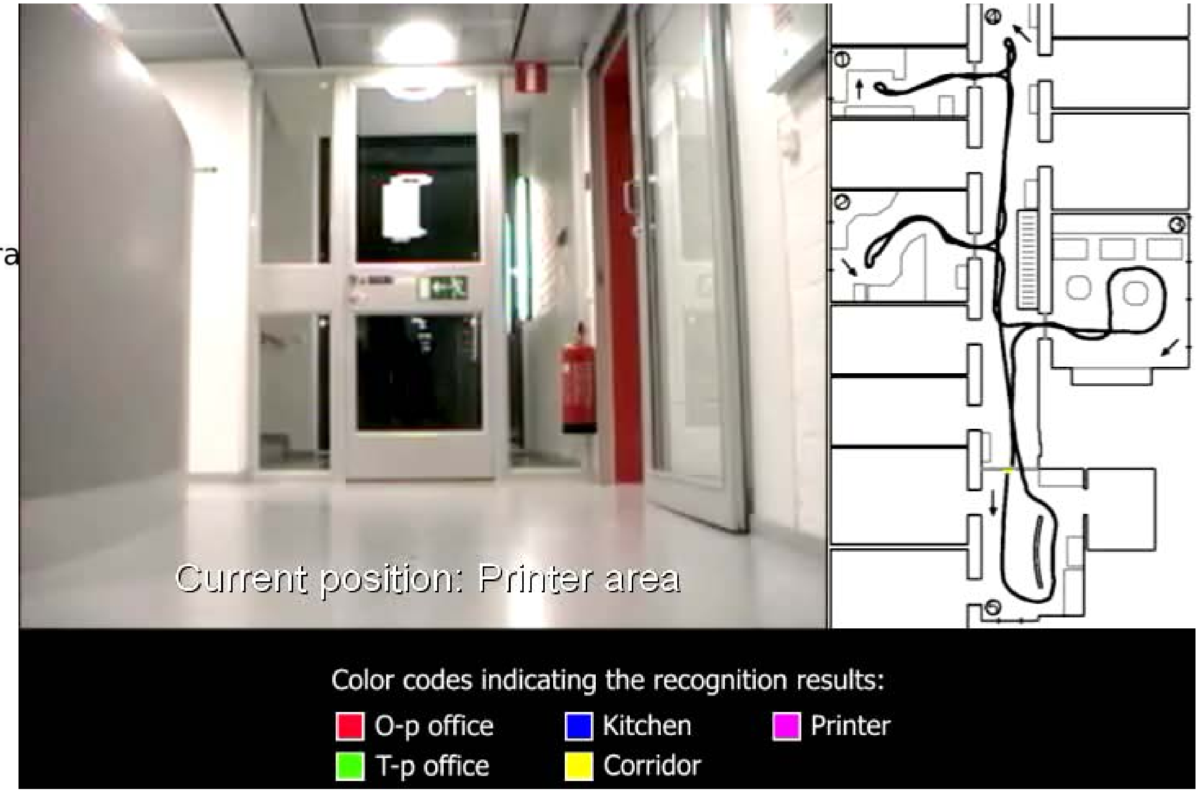
Goal

- Design a system able to visually categorize indoor places



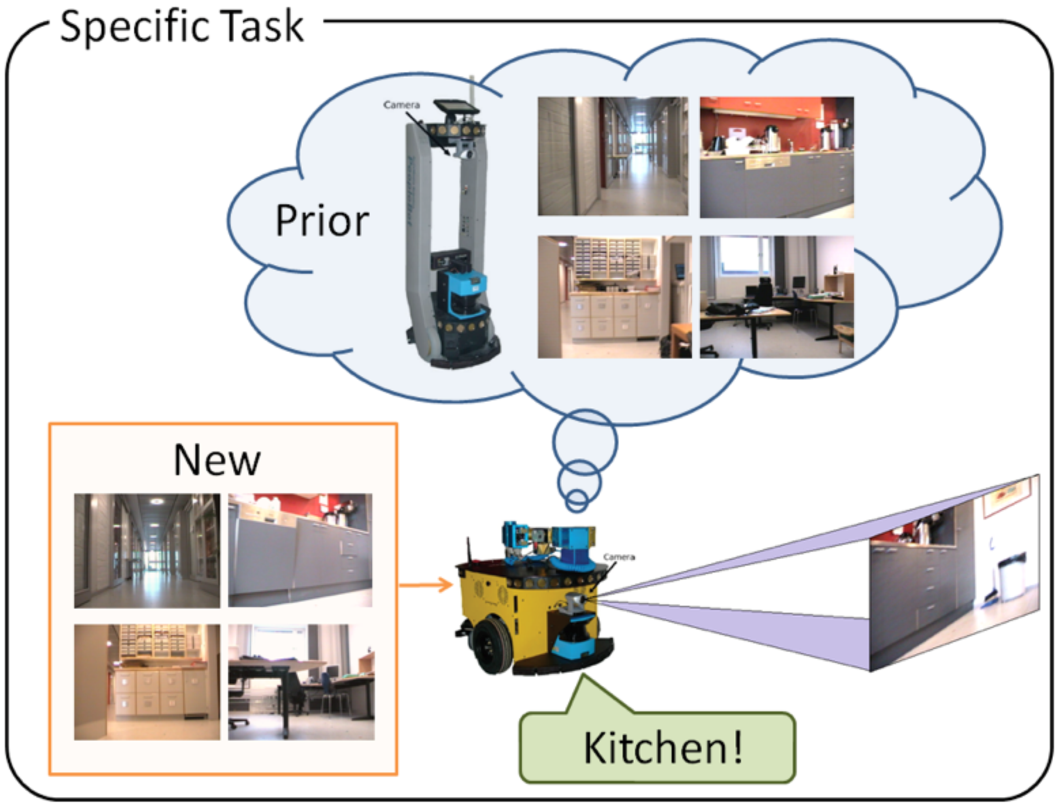
Why it is useful?

- Build a multi-layer representation of space and use it to navigate/interact with it

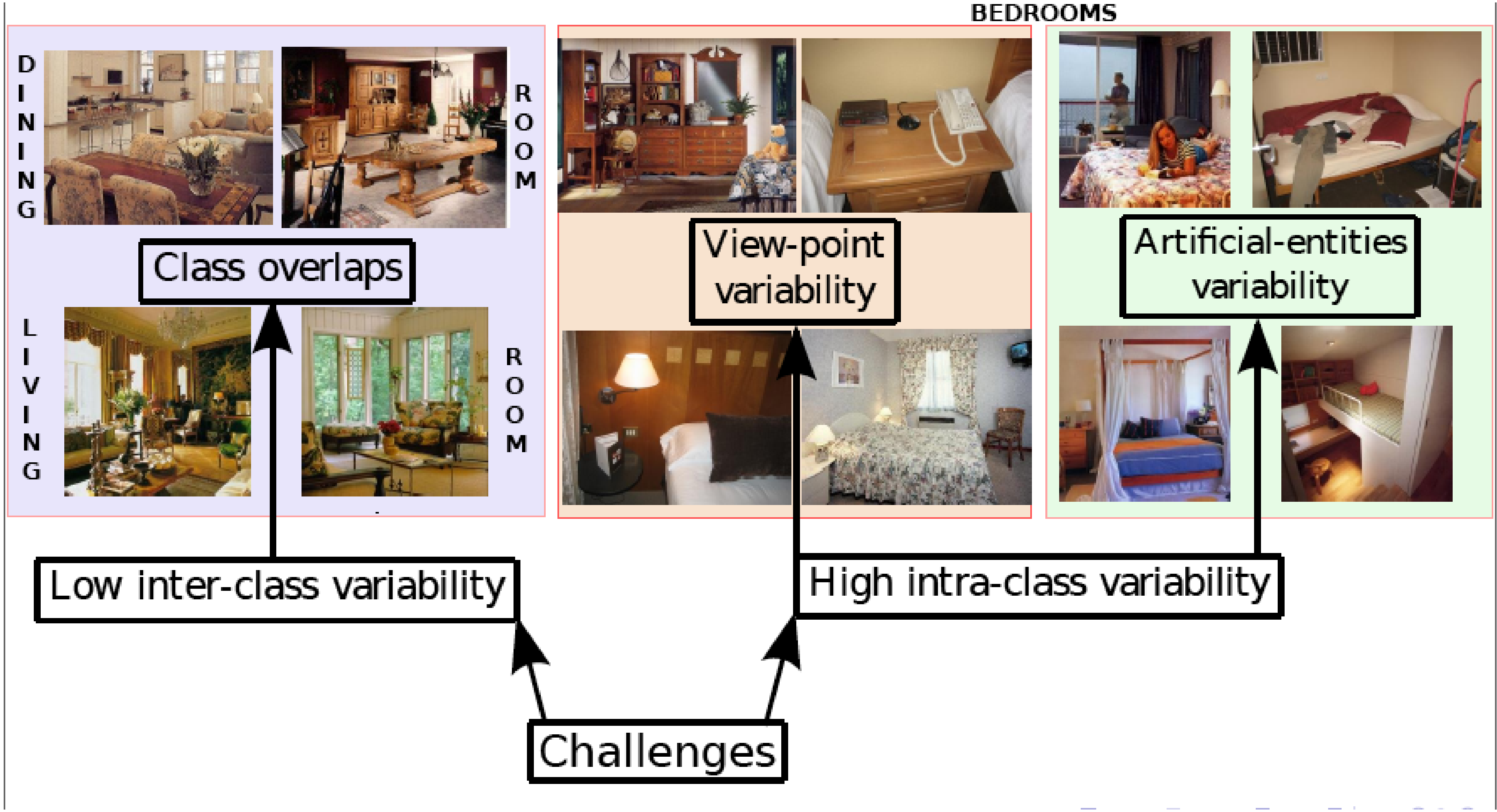


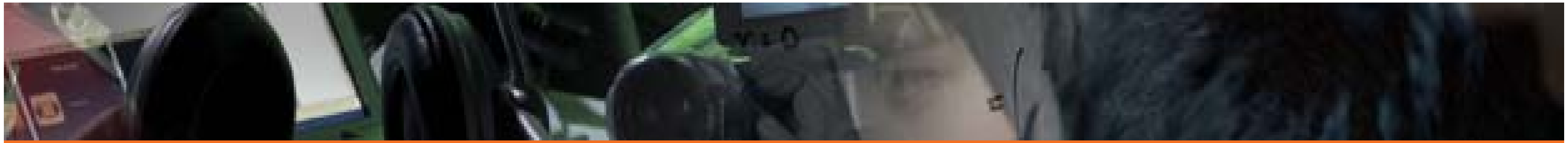
Why it is useful?

- Use categorical priors for boosting learning of new environments



Challenges


BEDROOMS
DINING
ROOM
Class overlaps
LIVING
ROOM
View-point
variability
Artificial-entities
variability
Low inter-class variability
High intra-class variability
Challenges



Human Visual Perception

- ***What do we remember and what do we forget when we recall a scene?***
  - WE DO REMEMBER: the gist of a scene, 4-5-landmark objects and their spatial configuration
  - WE DO **NOT** REMEMBER: all the objects in the scene, mid- to fine details

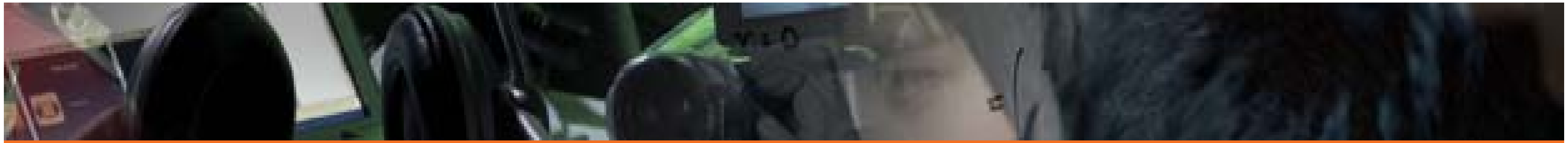
J. M. Wolfe. *Visual memory: what do you know about what you saw?* Current Biology, 1998, 8: R303-R304

Outline of the talk

- ***Focus on Gist-like representations***
  - *Visual Place Categorization*
  - *Transfer Learning of Visual Spatial Concepts*

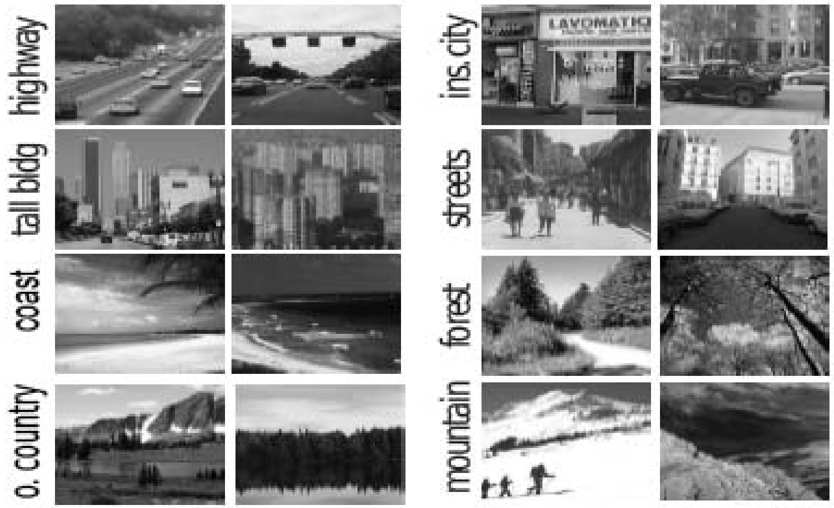
Outline of the talk

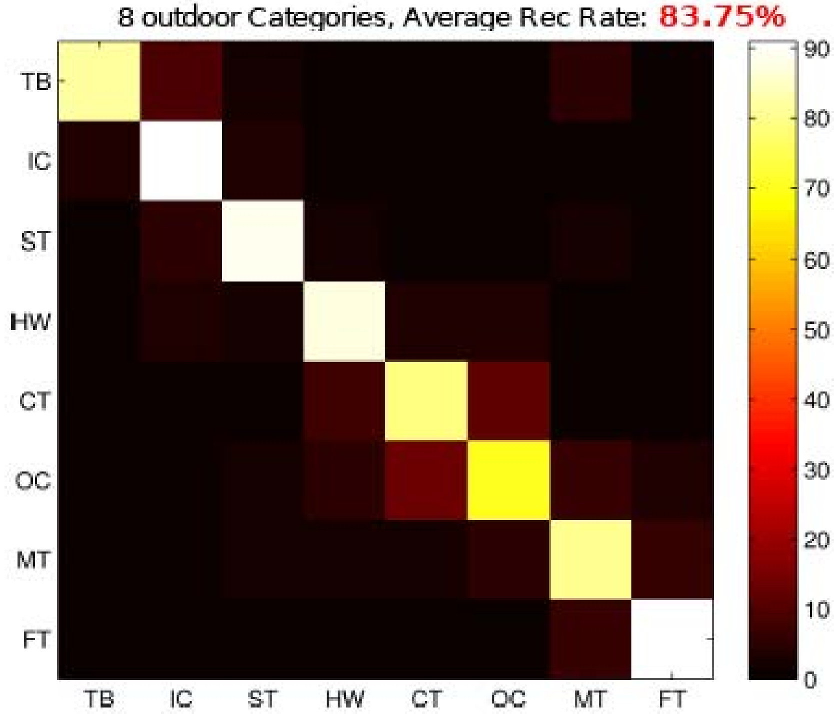
- ***Focus on Gist-like representations***
    - *Visual Place Categorization*
    - *Transfer Learning of Visual Spatial Concepts*



History: the CV Perspective

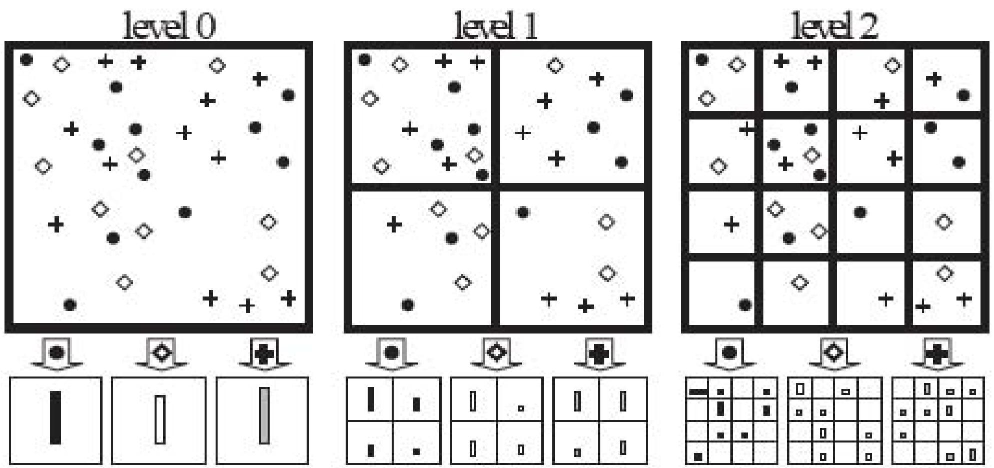
2001 A. Oliva and A. Torralba. Modeling the shape of the scene: A holistic representation of the spatial envelope. IJCV

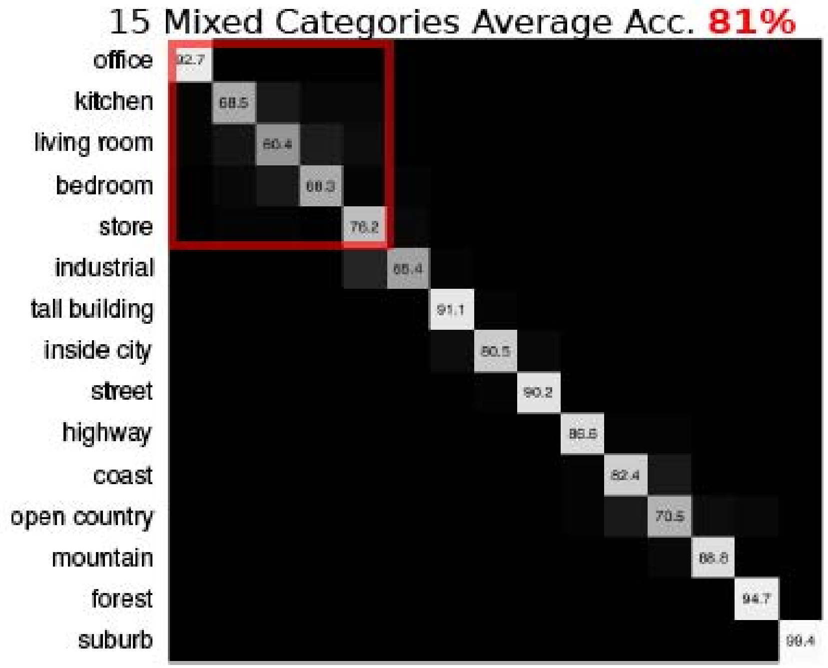




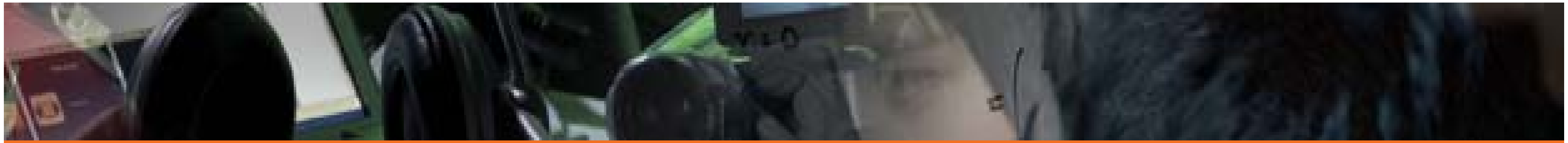
History: the CV Perspective

2006 S. Lazebnik, C. Schmid, and J. Ponce. Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. CVPR



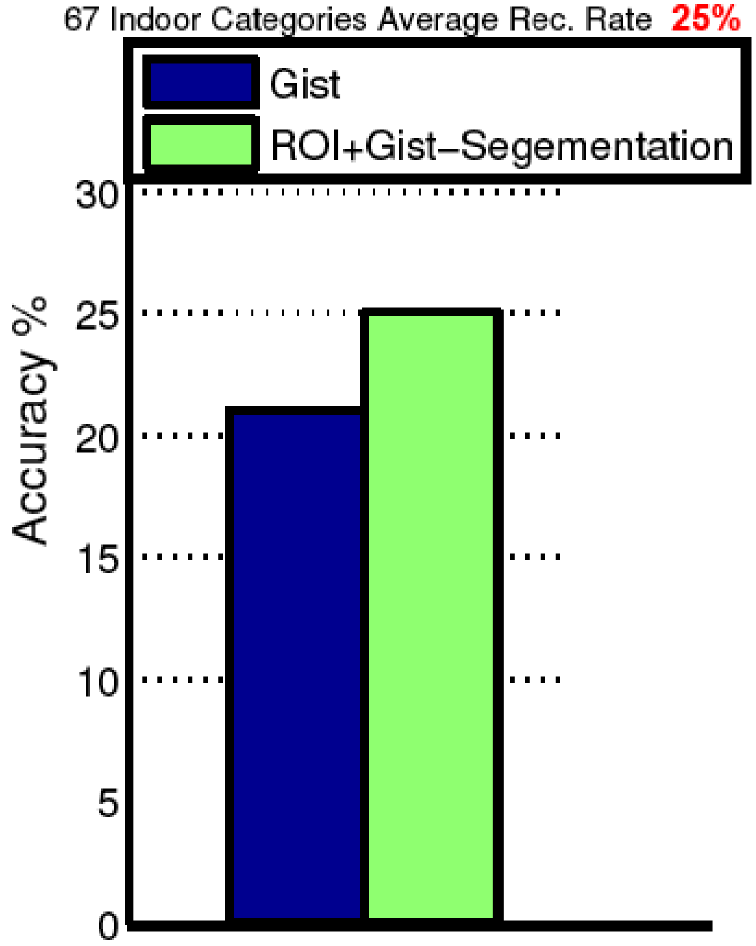


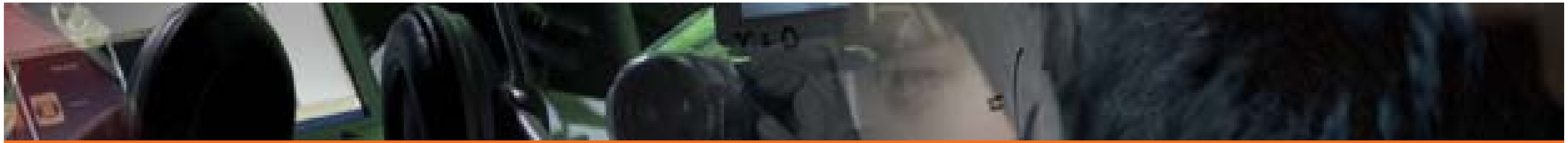
"Not surprisingly, confusion occurs between the indoor classes"



History: the CV Perspective

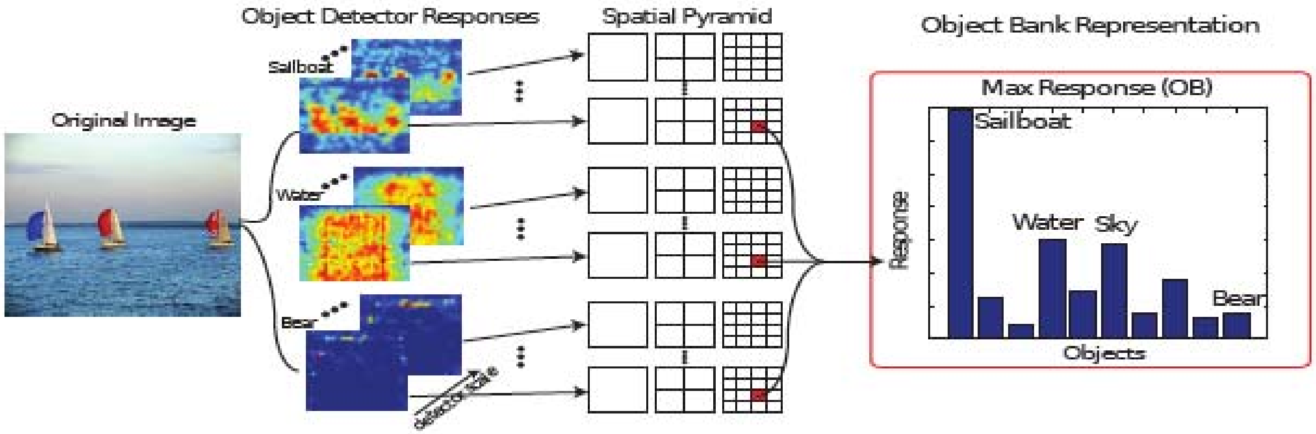
2009 A. Quattoni and A. Torralba. Recognizing Indoor Scenes. CVPR

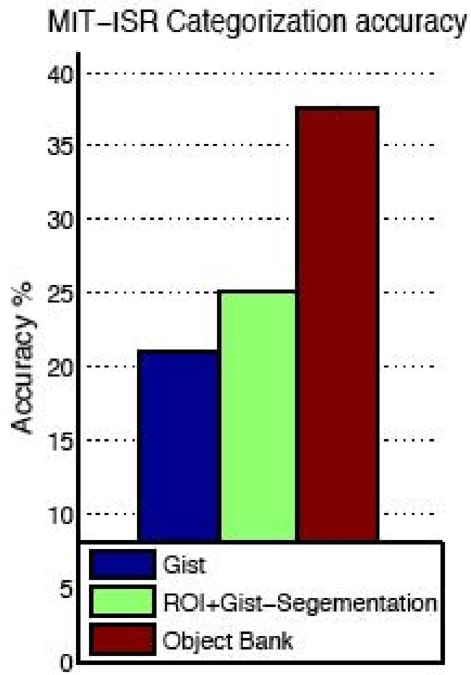




History: the CV Perspective

2010 Li-Jia Li et al. Object Bank: A High-Level Image Representation for Scene Classification and Semantic Feature Sparsification. NIPS

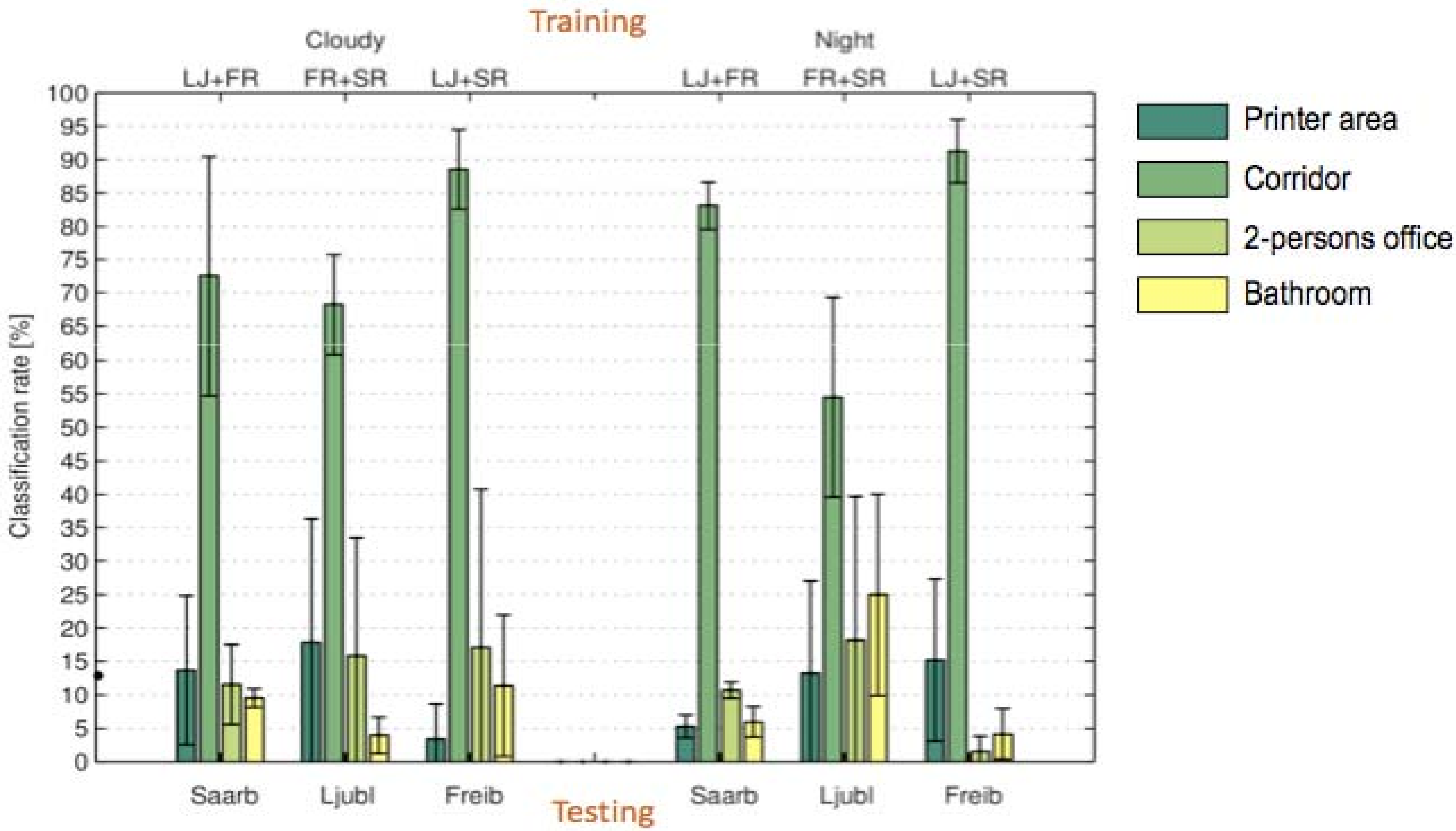




History: the Robotics Perspective

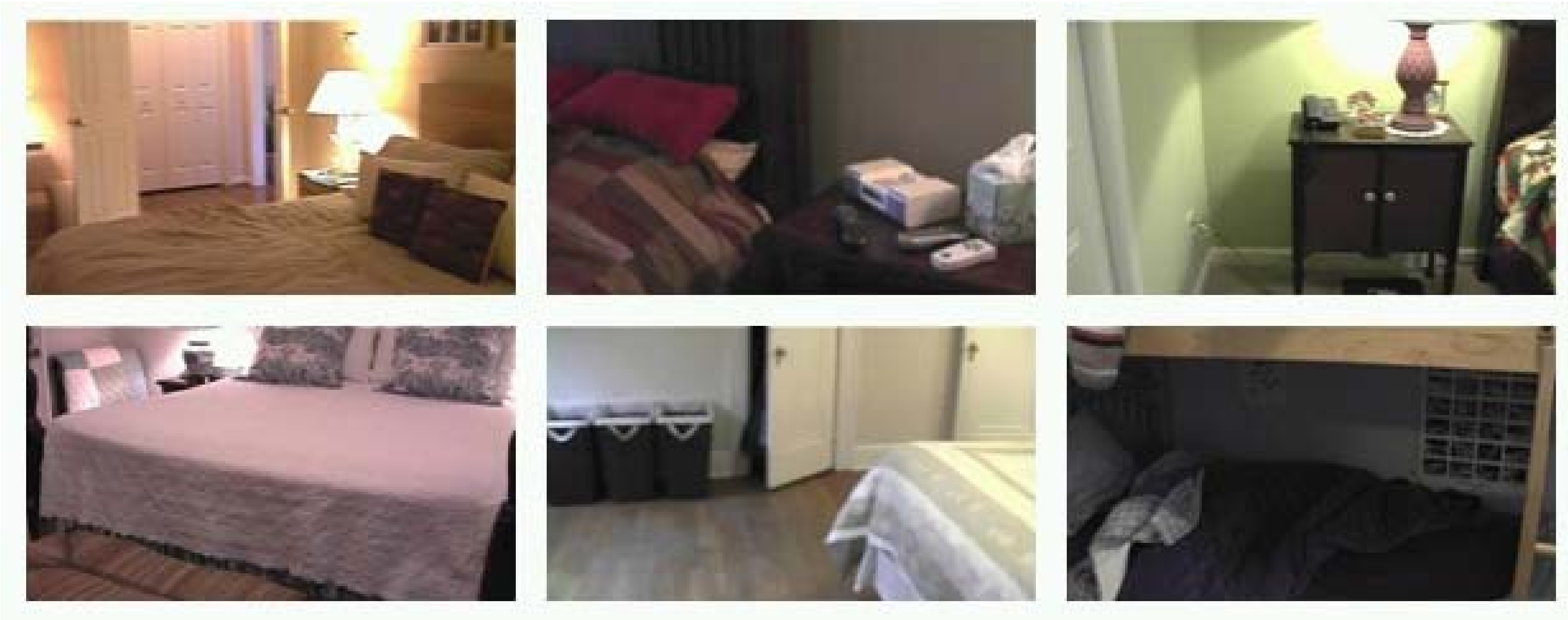
- A. Pronobis, B. Caputo. COLD: the Cosy Localization Database. IJRR 2009

Place Categorization



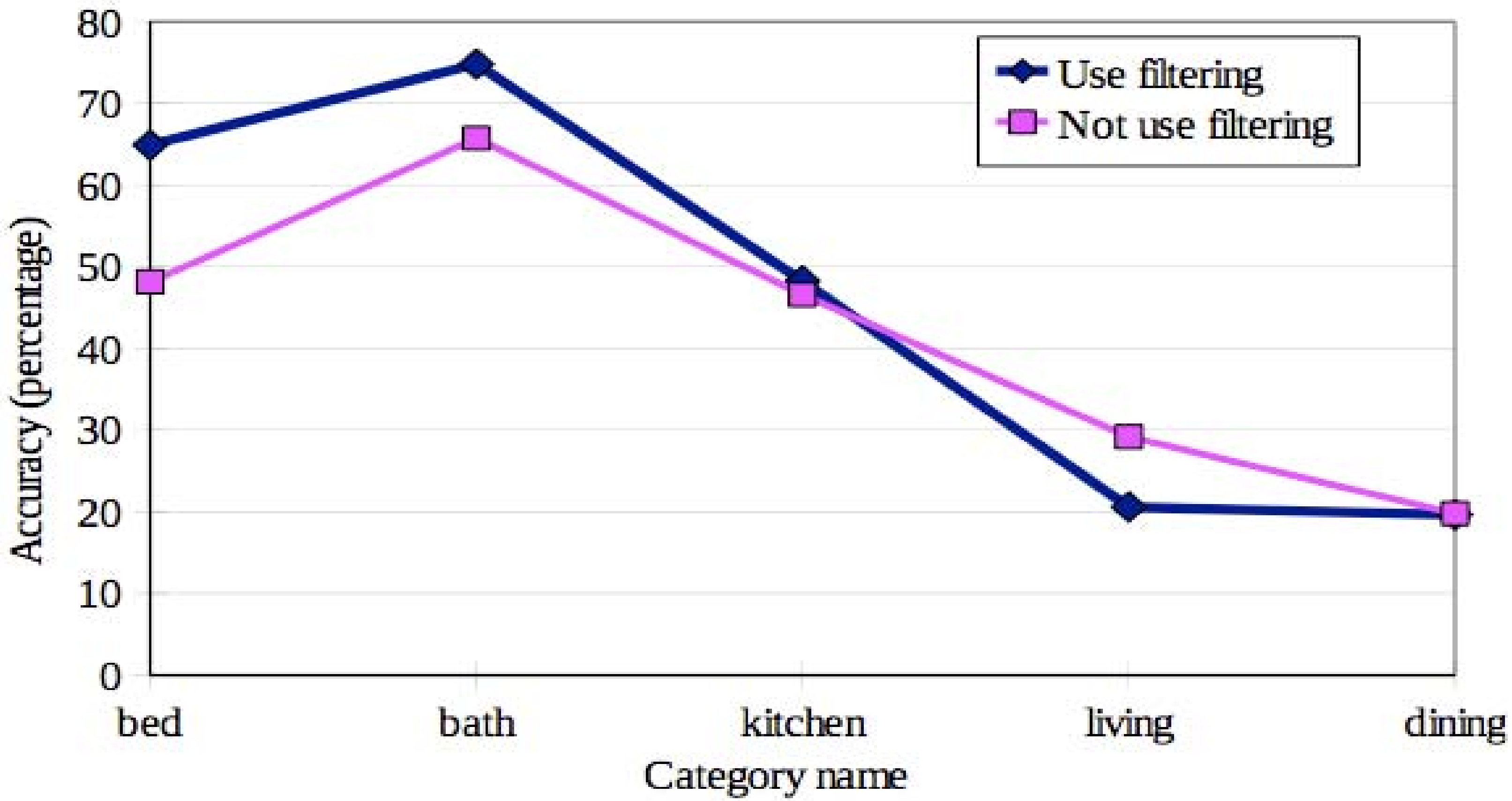
- Rec Rate SOTA: 27.57%

Place Categorization

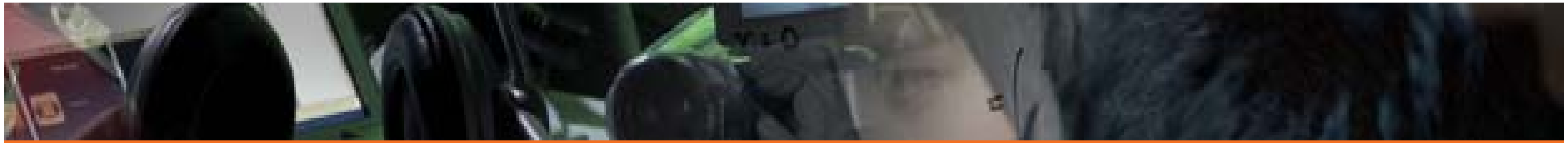


- J. Wu, H. Christensen, J. Rehg. Visual place categorization: problem, dataset, and algorithms.
- Proc of ICRA 2009.

Place Categorization



- Rec Rate SOTA: 45.62%

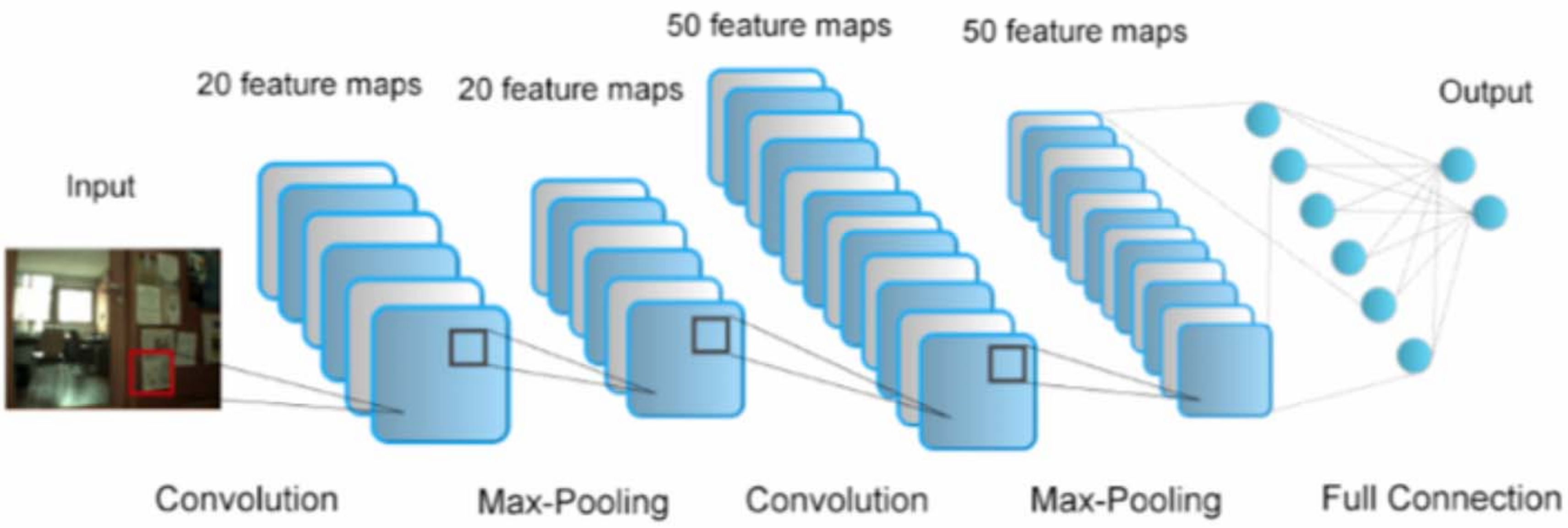


Place Categorization

Gist representations
do not work……
or do they?

Place Categorization

M. Mirza, M. Fornoni, R. Collobert, B. Caputo. *On Learning Visual Features for Indoor Scenes.* ICRA12, submitted



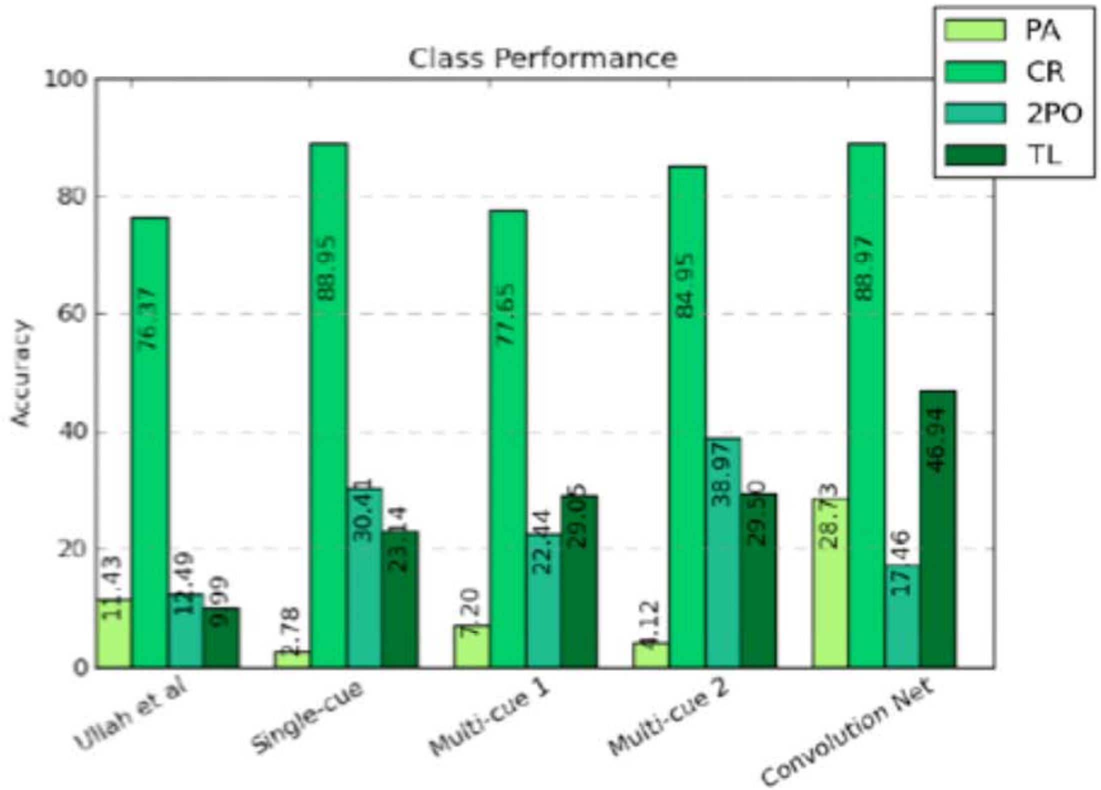
We learn the gist as opposed to designing it!

Place Categorization

- We follow the approach outlined in *[Jarret et al, ICCV09]* for object recognition benchmarks
- ***We use a two stages architecture for feature extraction, with each stage composed by:***
  - *20/50 filter banks of size (10x10), (4x4)/(2x2) shaped max pooling, rectification and non-linearity*
- ***Third stage:*** *hidden layer with 1,000 nodes, connected to an output layer of the same dim of the number of classes*

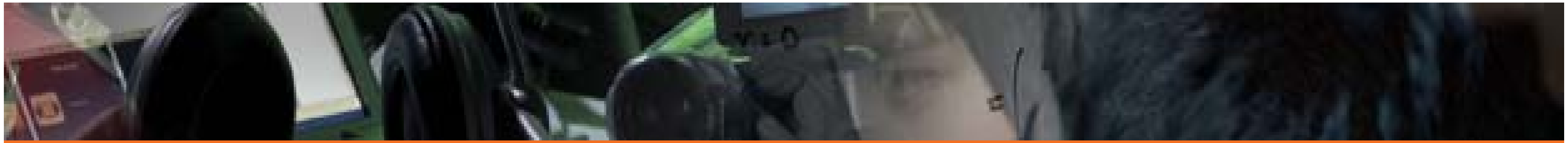
Place Categorization

Results: COLD database



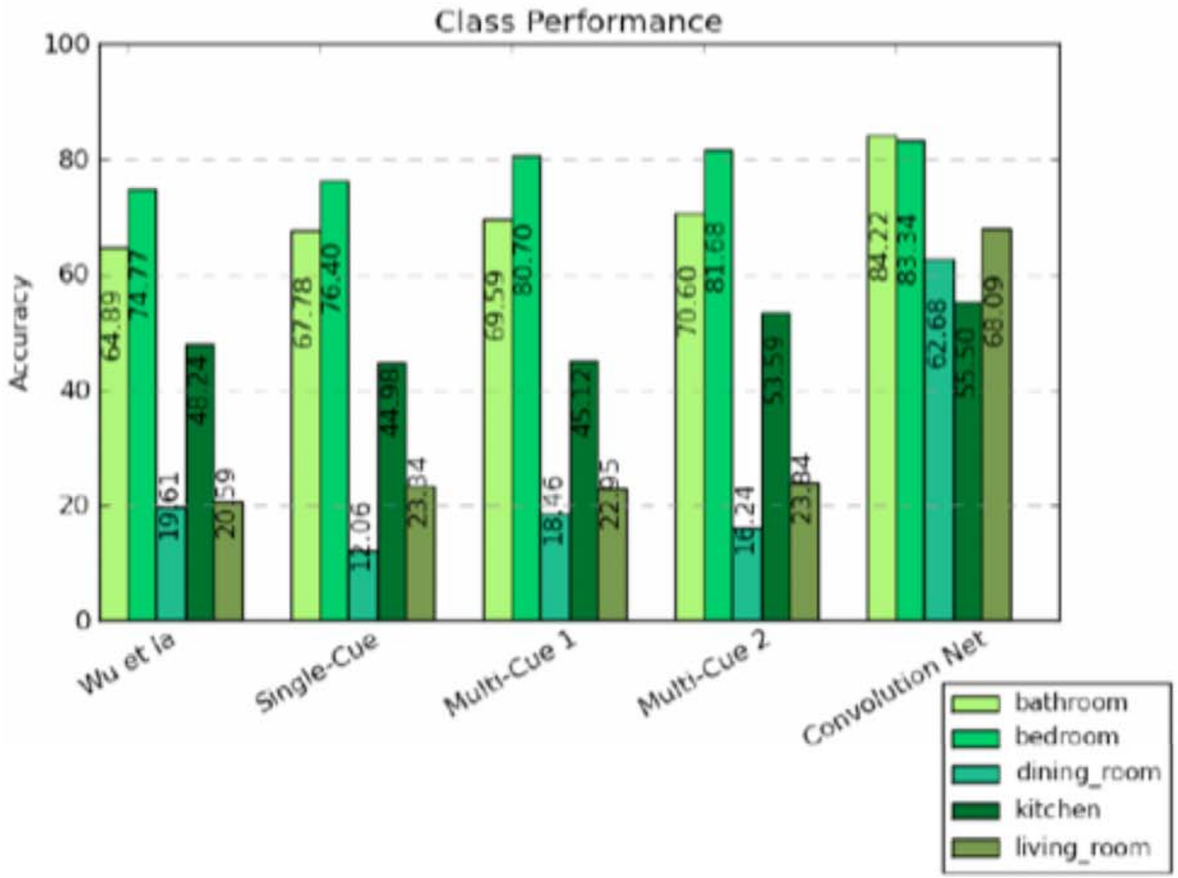
rec rate sota: 27.57%

rec rate dcn: 45.52%



Place Categorization

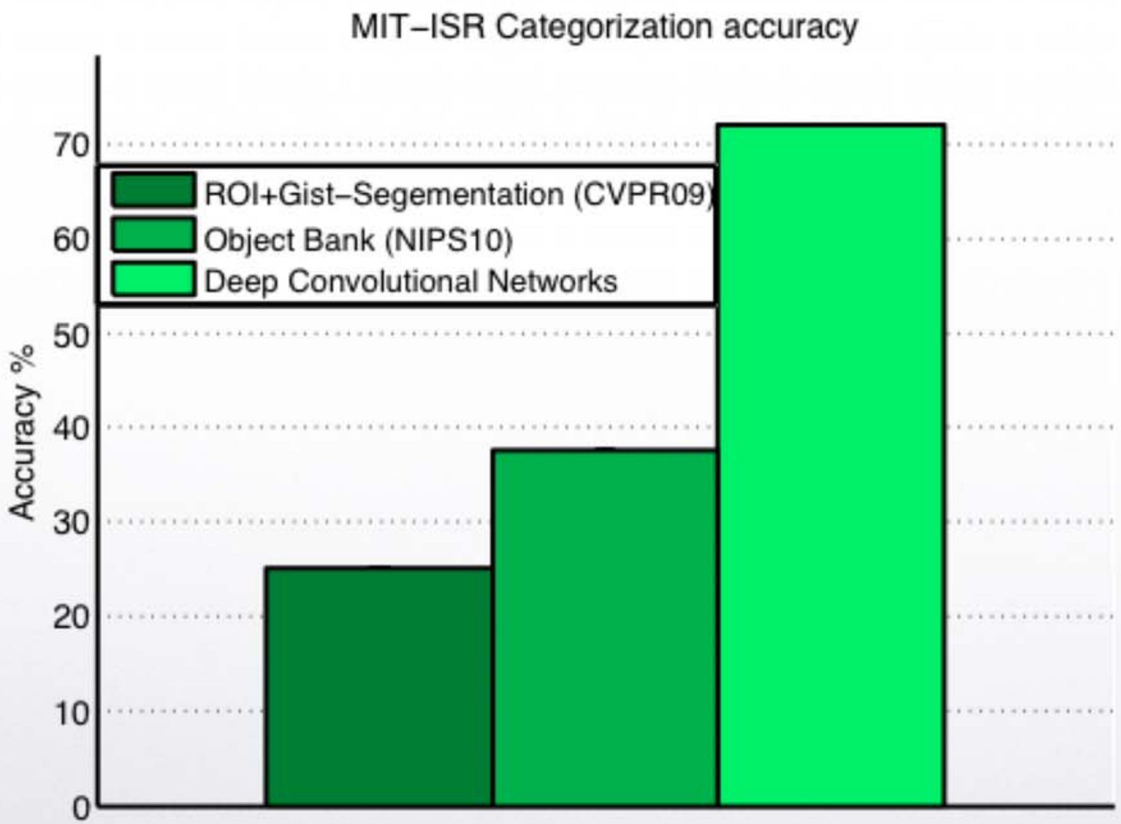
Results:VPC database



rec rate sota:45.62%

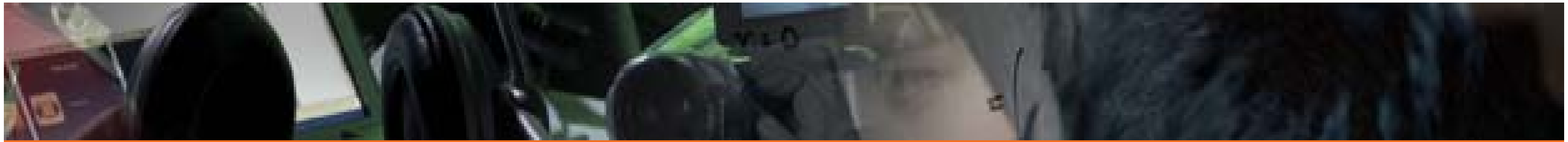
rec rate dcn: 70.76%

Place Categorization



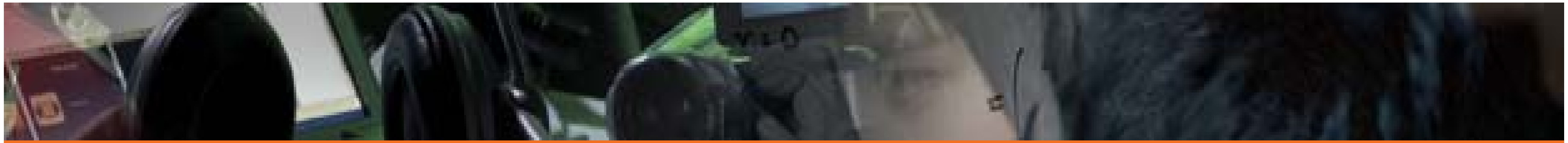
rec rate sota:37%

rec rate dcn: 72%



Outline

- ***Focus on Gist-like representations***
    - *Visual Place Categorization*
    - *Transfer Learning of Visual Spatial Concepts*



Transfer Learning

Learning Categories from few examples

- Basic Intuition

  IF a system has already learnt k categories

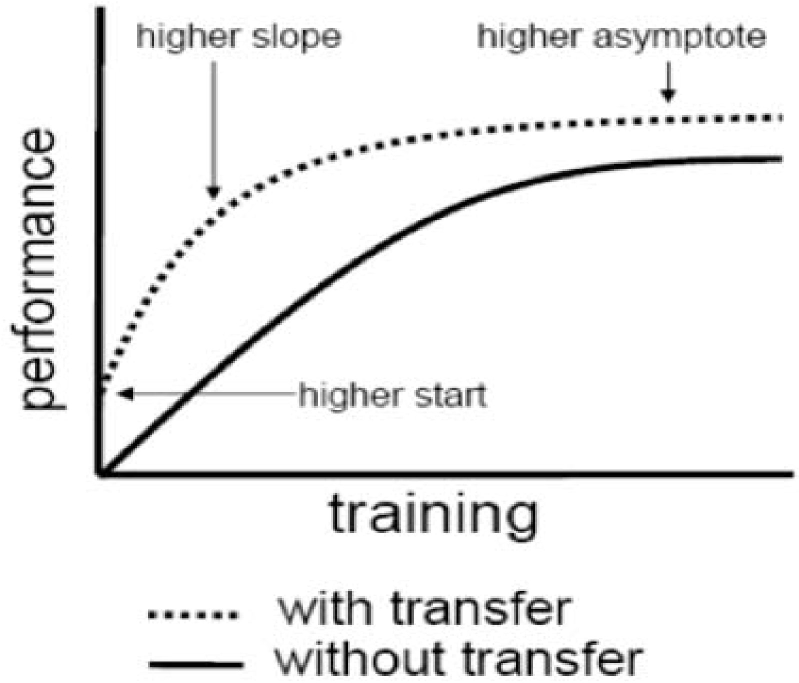
  THEN learning the $(k+1)^{th}$ should be easier.

Transfer information from prior knowledge

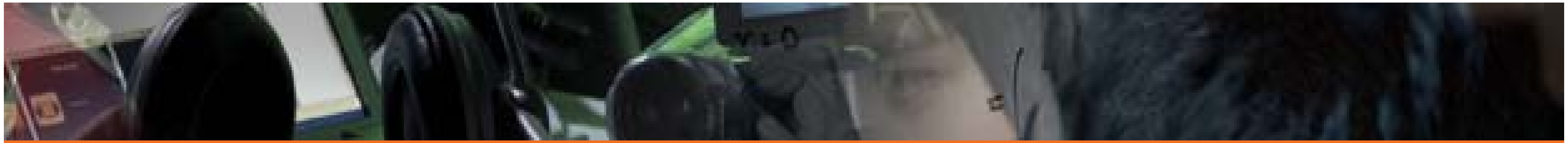
- **Hypothesis**: the more familiar the new category is to prior knowledge, the easier the learning process is.

Transfer Learning

Advantages



- + performance higher than that of an ignorant agent;
- + less time needed than learning from scratch;
- + heigher asymptotic performance than using only train data.



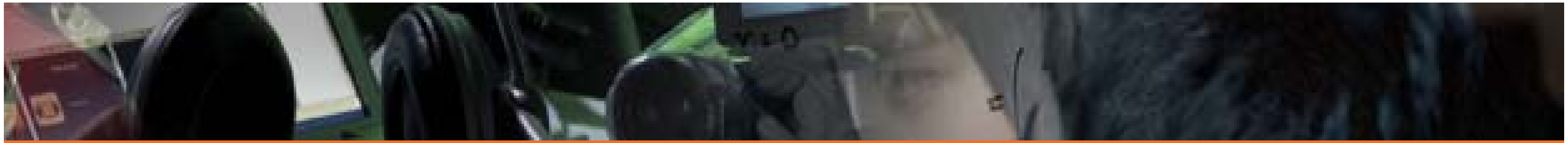
Transfer Learning

Transfer Learning: key questions

- **How to transfer**: using Least-Square Support Vector Machine.

- **What to transfer**: we define the prior knowledge as the hyperplanes of the classifiers $\mathbf{w}'_j$, j=1…k of the k classes already learned.

- **When/How much to transfer**: using the Leave One Out error on the training set. It is an unbiased estimator of the classifier generalization error.



Transfer Learning

Transfer Learning from Multiple Classes

Weighted LS-SVM + Adaptation

$$\min_{\mathbf{w},b} \frac{1}{2}\left\| \mathbf{w} - \sum_{j=1}^{k} \beta_j \mathbf{w}'_j \right\|^2 + \frac{C}{2}\sum_{i=1}^{n} \zeta_i$$

Leave-One-Out Error : $r_i^{(-i)} = f(\mathbf{w}'_{1\ldots j}, b_{1\ldots j})$

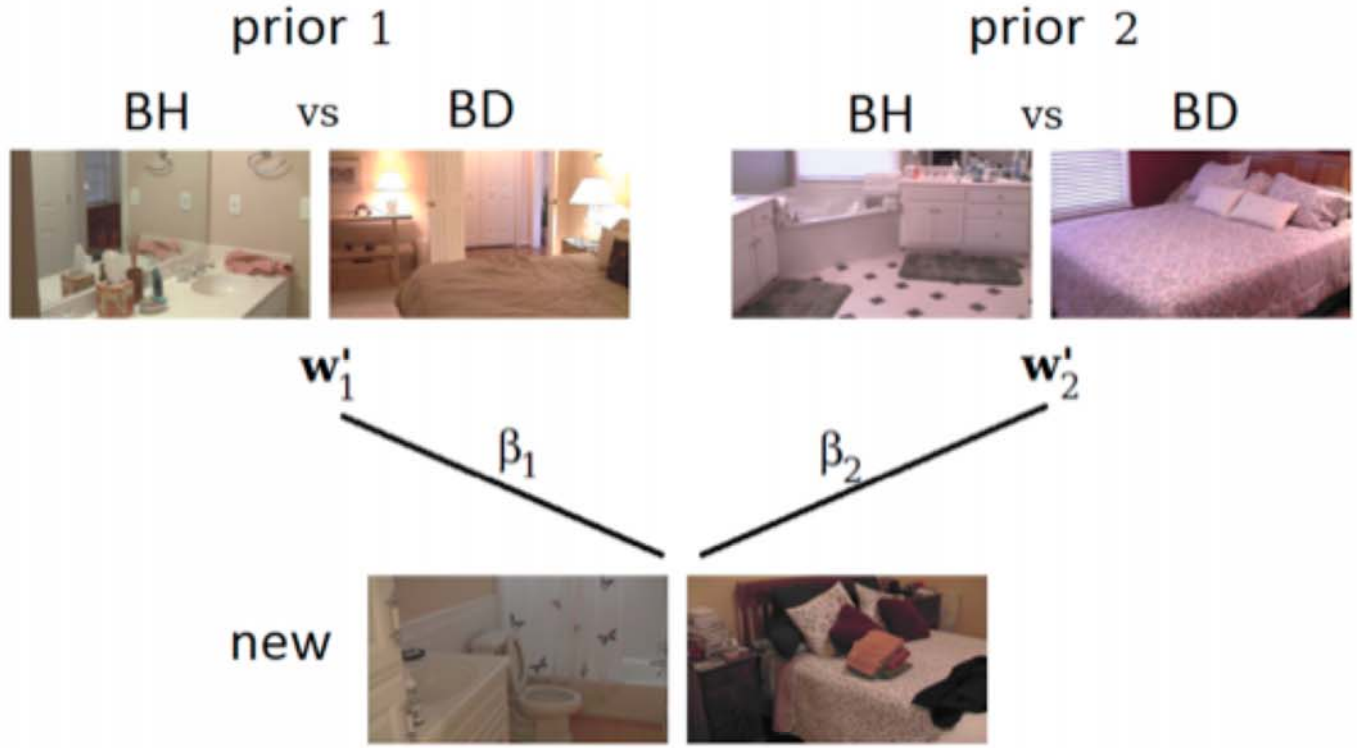
Leave-One-Out Prediction: $\mathbf{y} = g(\mathbf{r}_i^{(-i)})$

Loss Function: $l(y_i, \tilde{y}_i) = \zeta_i \max\left[1 - y_i\tilde{y}_i, 0\right]$

Objective Function: $J = \sum_{i=1}^{n} l(y_i, \tilde{y}_i) \quad \text{s.t.} \quad \|\boldsymbol{\beta}\|_2 \leq 1$

Find the best b_j minimizing J

Transfer Learning



$$\min_{\mathbf{w}} \| \mathbf{w} - (\beta_1 \mathbf{w}'_1 + \beta_2 \mathbf{w}'_2) \|^2$$

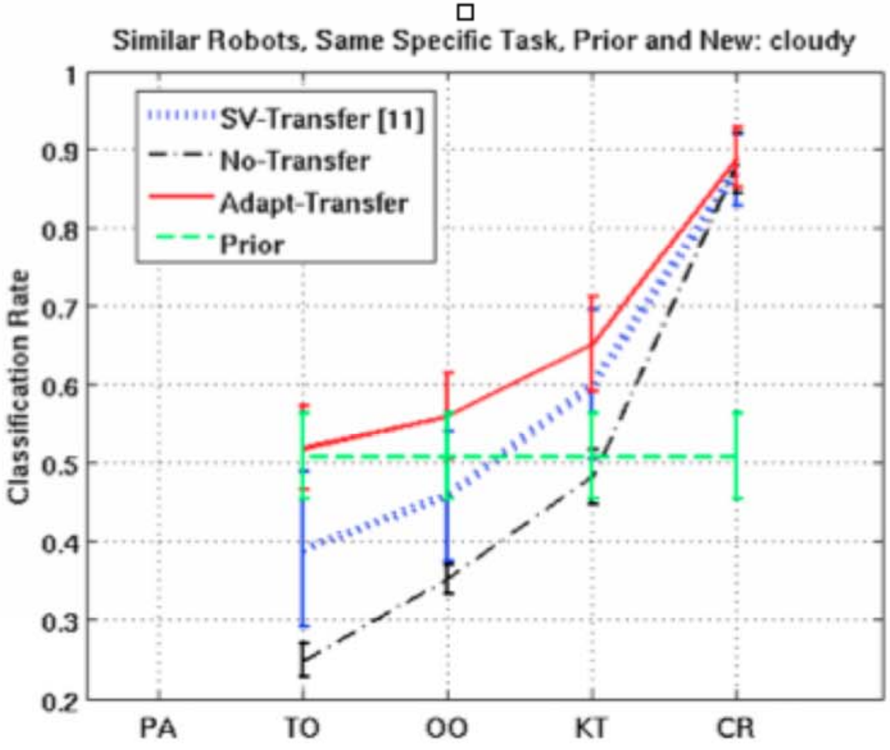
T. Tommasi, F. Orabona, B. Caputo. *Safety in numbers: learning categories from few examples with multi model knowledge transfer* . Proc CVPR 2010

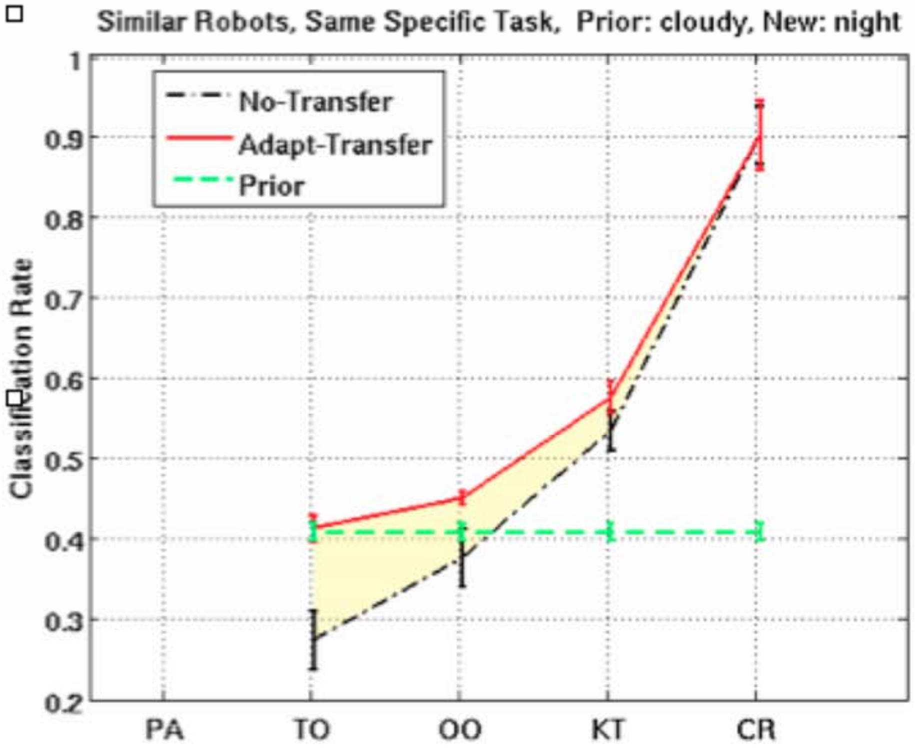
Transfer Learning

S. P. Elango, T. Tommasi, B. Caputo. *Transfer learning of visual concepts across robots: a discriminative approach*. IEEE T on Robotics, under review, 2011

IDOL DB

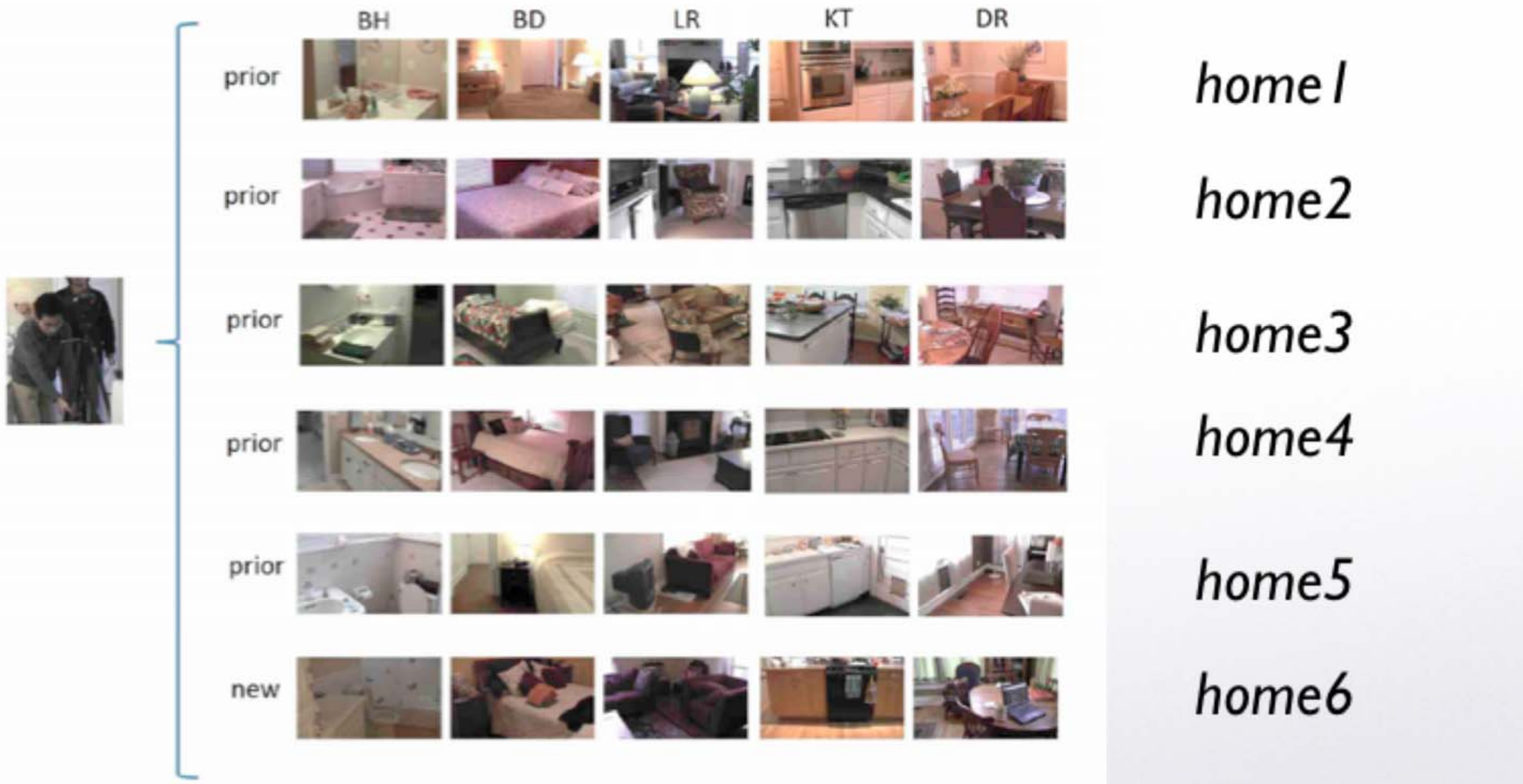
Results: IDOL database





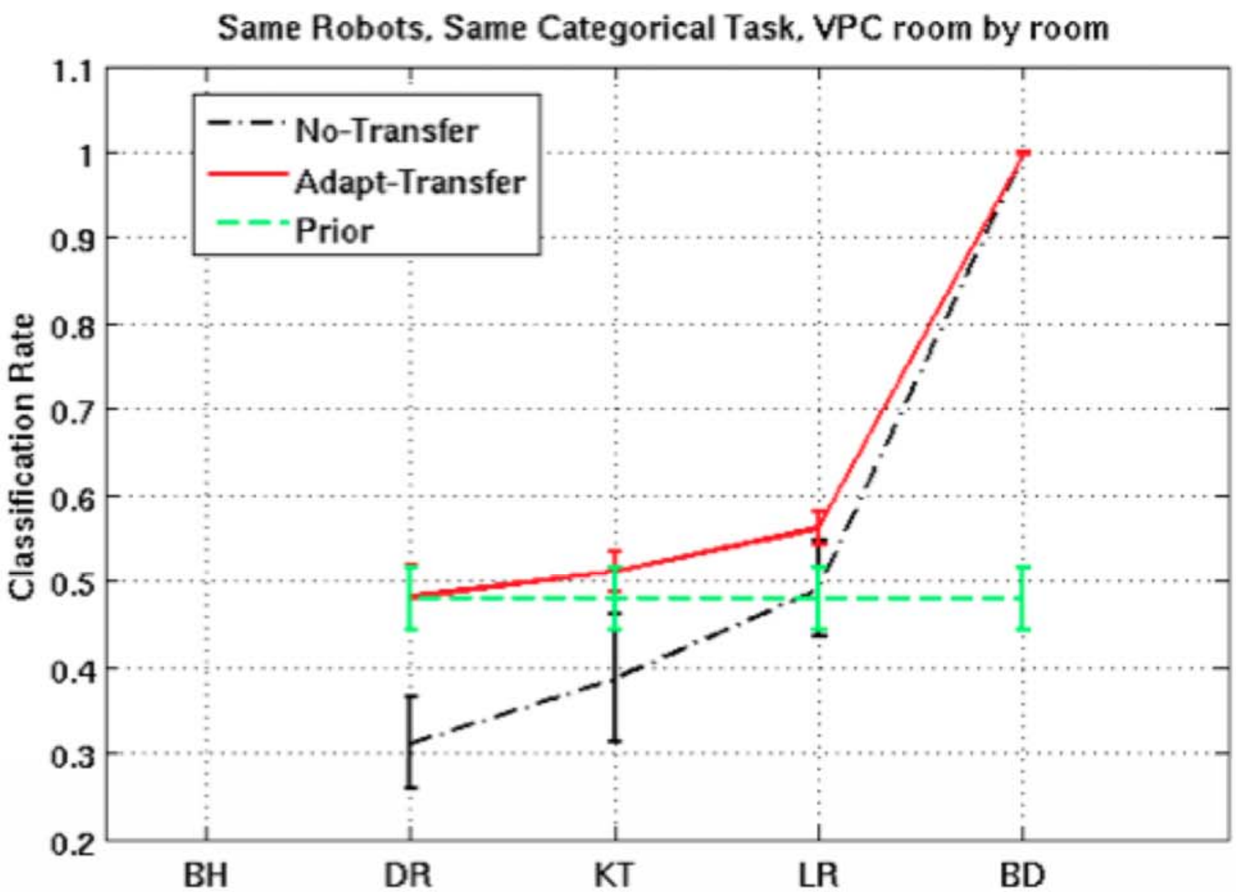
Transfer Learning

VPC database



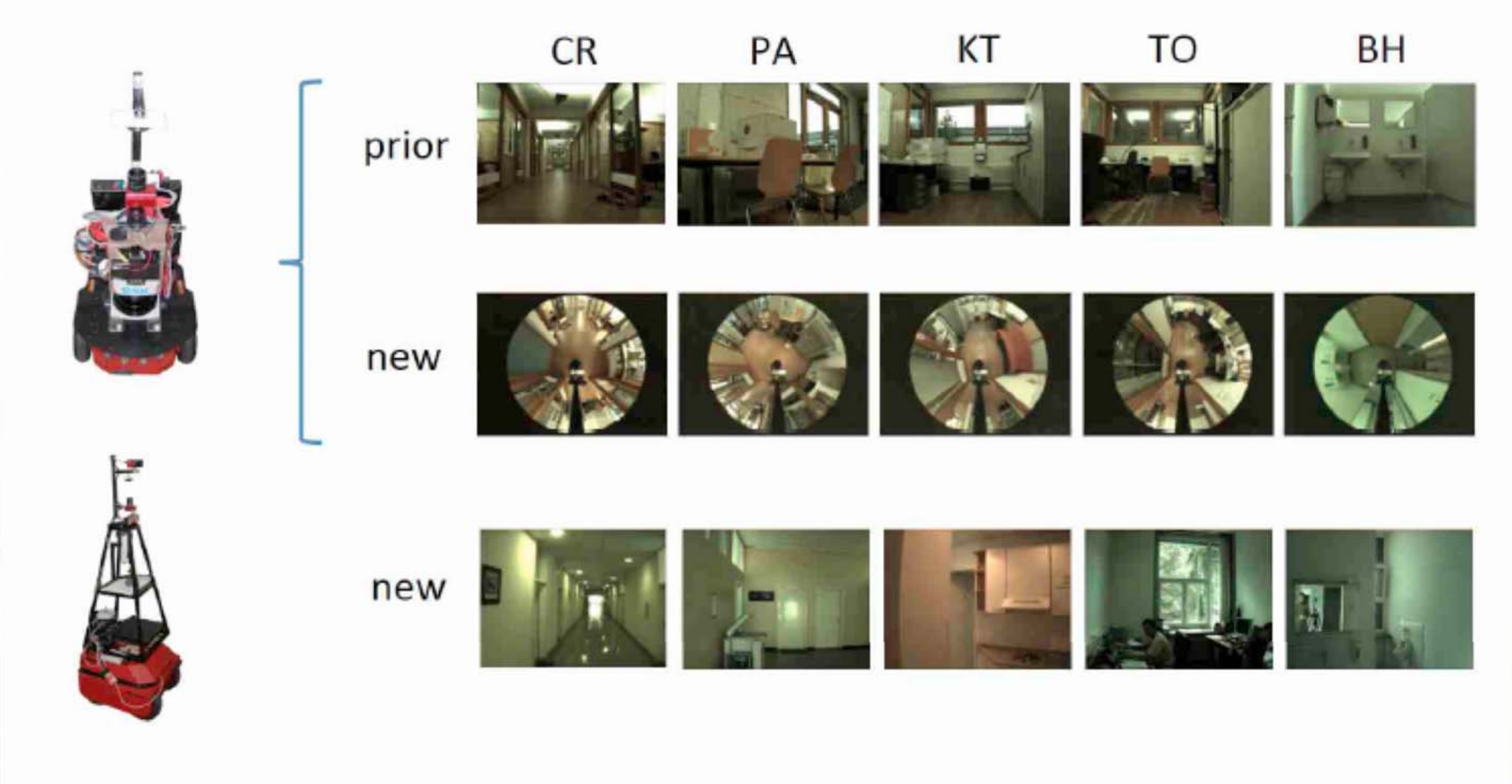
Transfer Learning

Results: VPC database

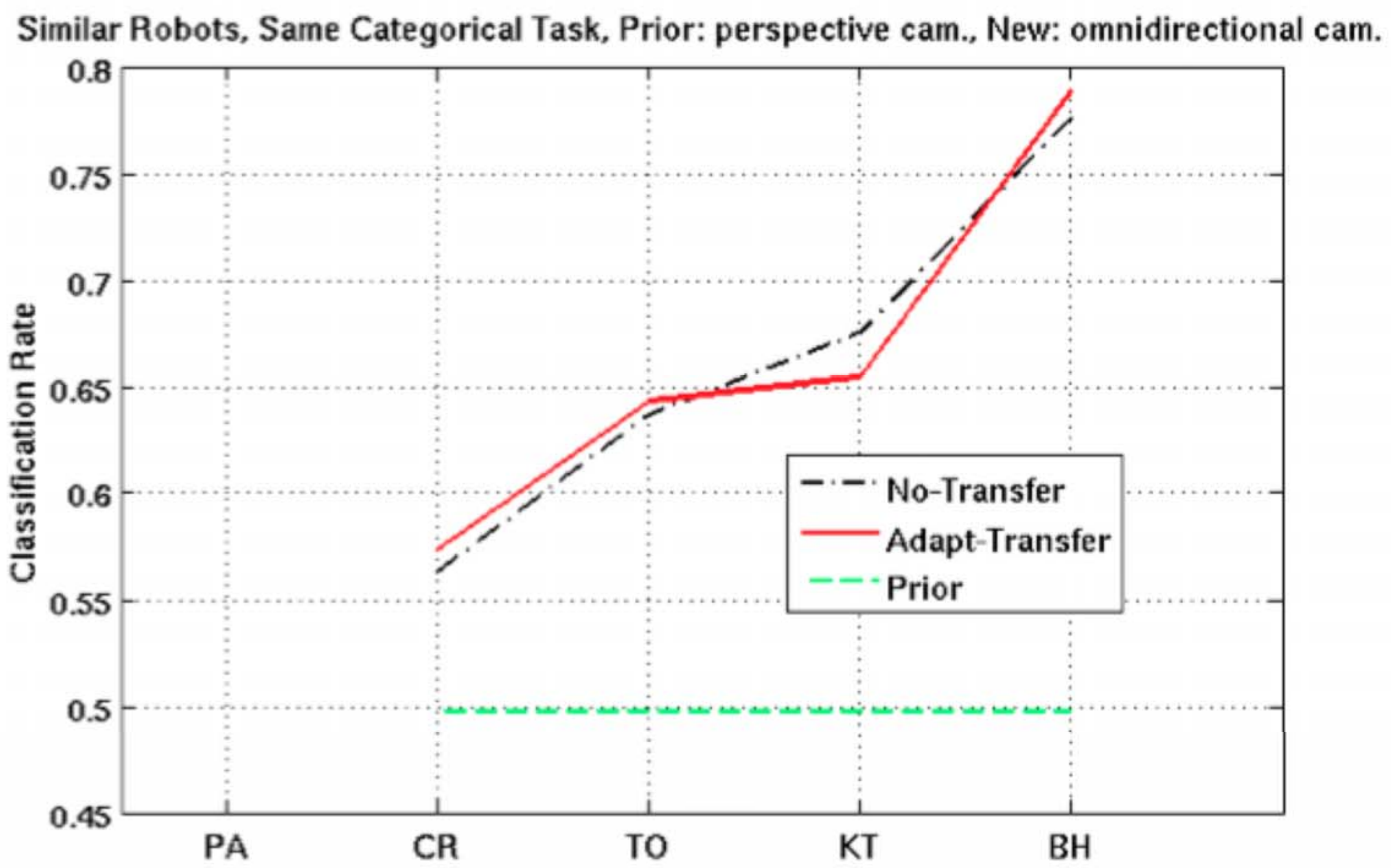


Transfer Learning

COLD database



Results: COLD database



Wrapping up

- ***Focus on Gist-like representations***
  - *Visual Place Categorization: shows big promise by learning the gist representation from the data*
  - *Transfer Learning of Visual Spatial Concepts: works as long as the prior is informative*

Acknowledgments

- DIRAC EU Integrated Project (FP6)
  - Work on transfer learning (T. Tommasi, F. Orabona)
- ICS-SNSF
  - Work on learning perceptual features (M. Fornoni. M. Mizra)
- SS2-Rob, EU NoE PASCAL Pump Priming
  - Work on transfer learning across heterogeneous robot platforms (S. P. Elango)