

# Using dependent patches for object categorization in a generative framework



Gang Wang<sup>1,2</sup>

Ye Zhang<sup>2</sup>

Li Fei-Fei<sup>1</sup>

<sup>1</sup> University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

<sup>2</sup> Harbin Institute of Technology, Harbin, China

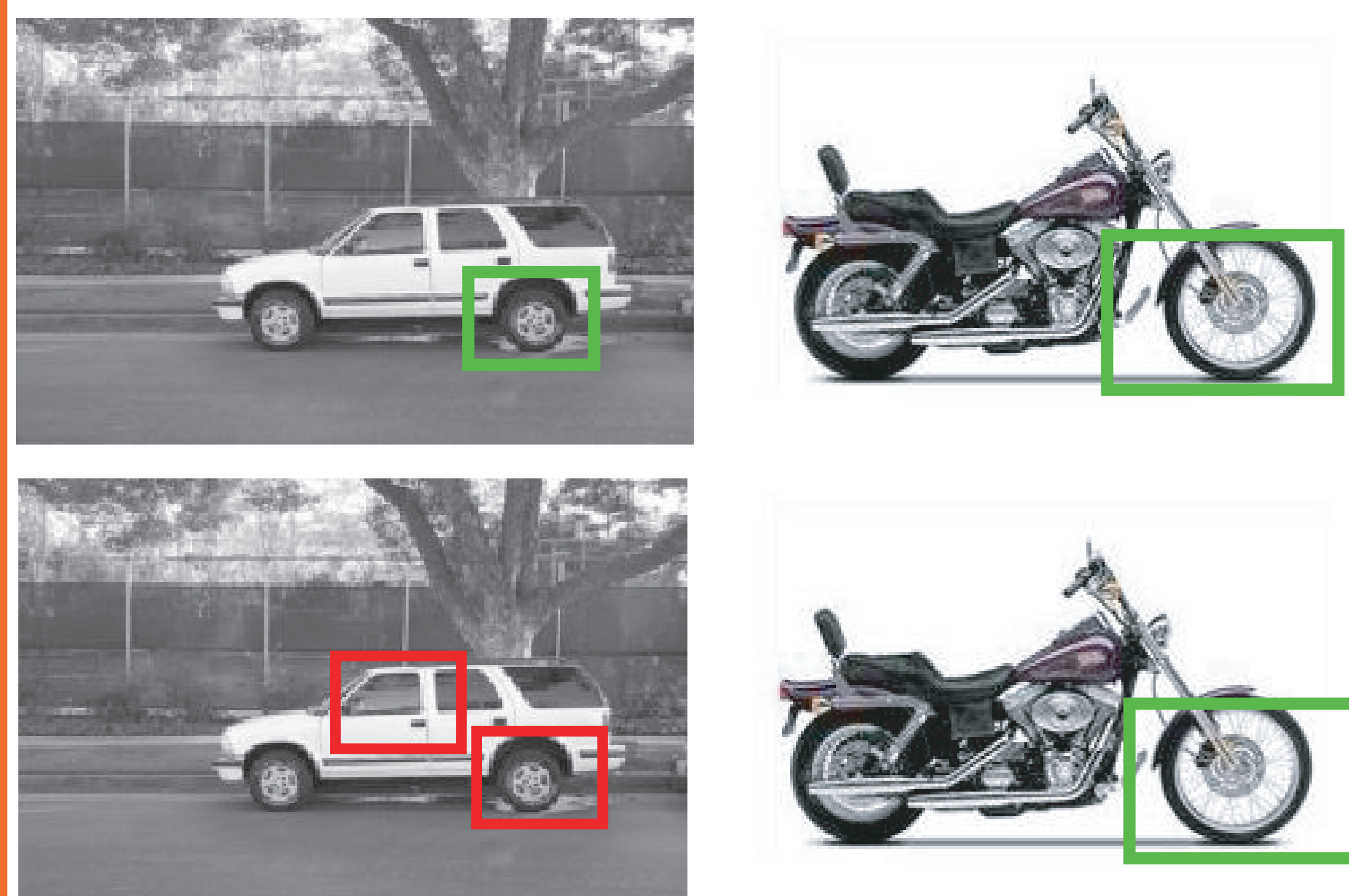
## 1. Summary

**Problem statement:** Categorizing objects by modeling the inter-dependency of local patches.

### Contributions:

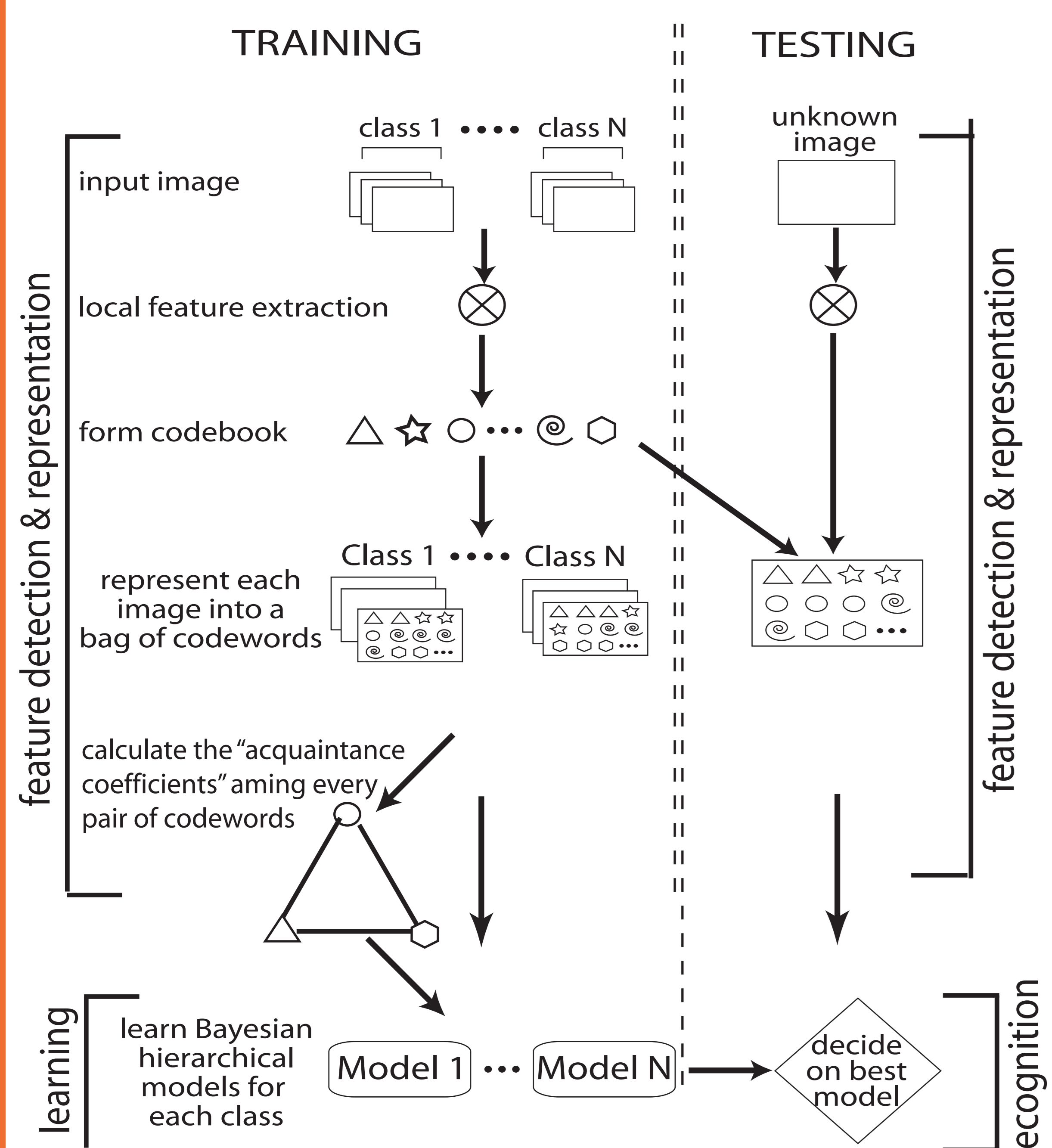
1. We propose a Bayesian model that relaxes the independence assumption of "bag of words" model and can exploit patch dependency.
2. We get a best recognition performance in both Caltech 4 dataset and Caltech 101 dataset.
3. We extract the first sensible taxonomy on the Caltech 101 dataset.

## 2. Motivation

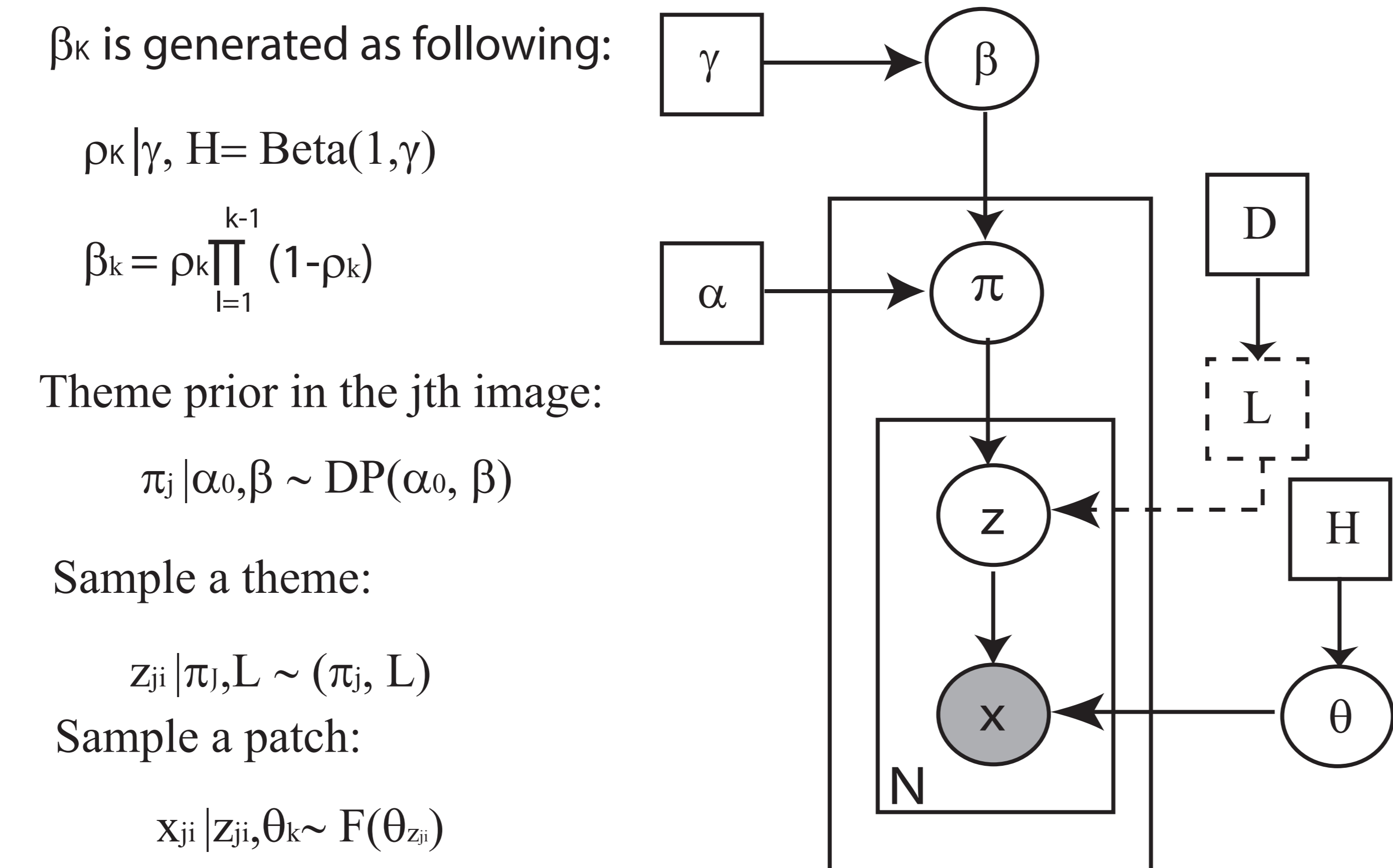


Patch dependency encodes much useful information for object categorization.

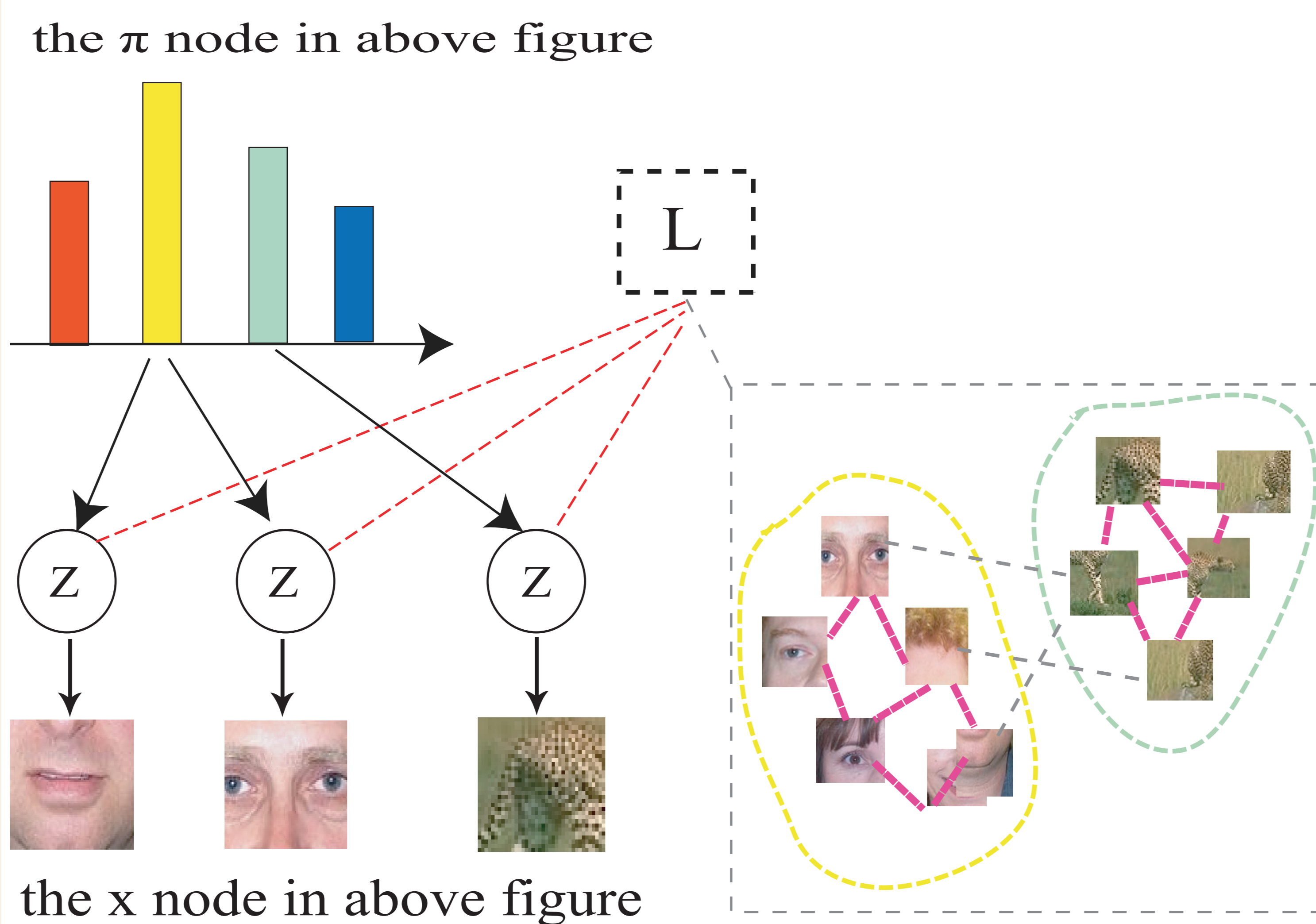
## 3. Algorithm



## 4. The Generative Model

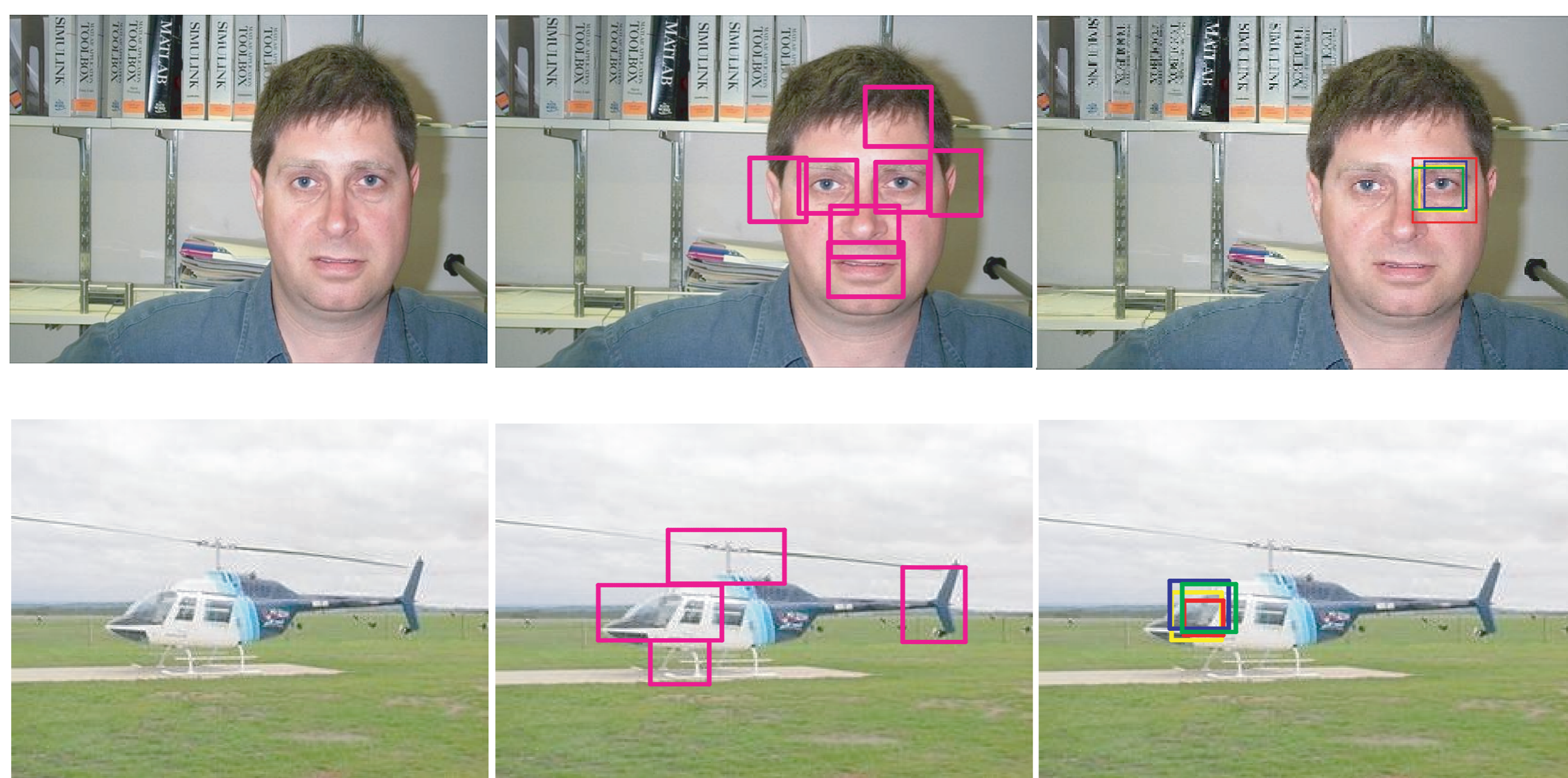


## 5. A "Zoom in" version



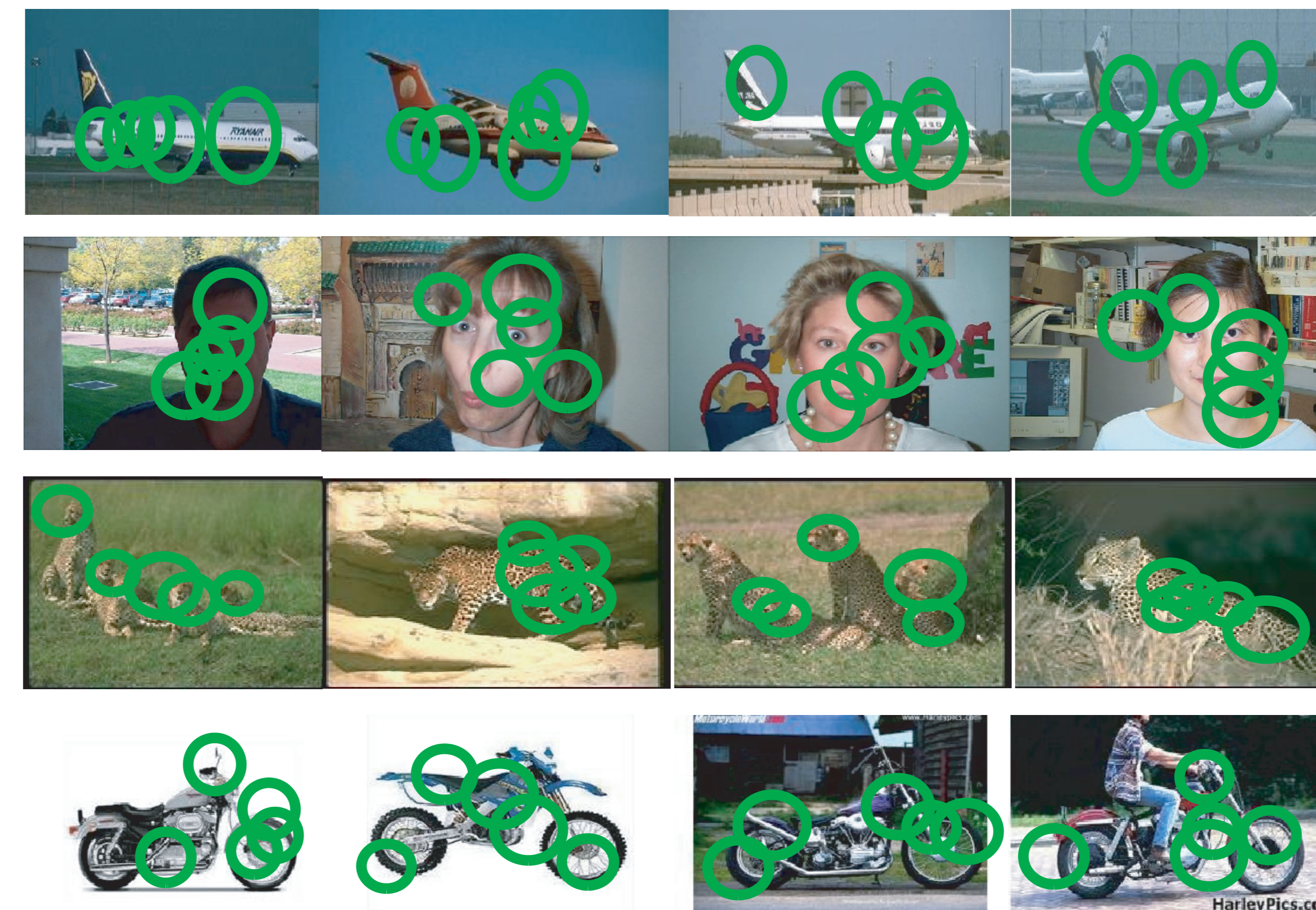
## 6. Semi-supervised variation

Label distinctive parts and assign themes for theme.

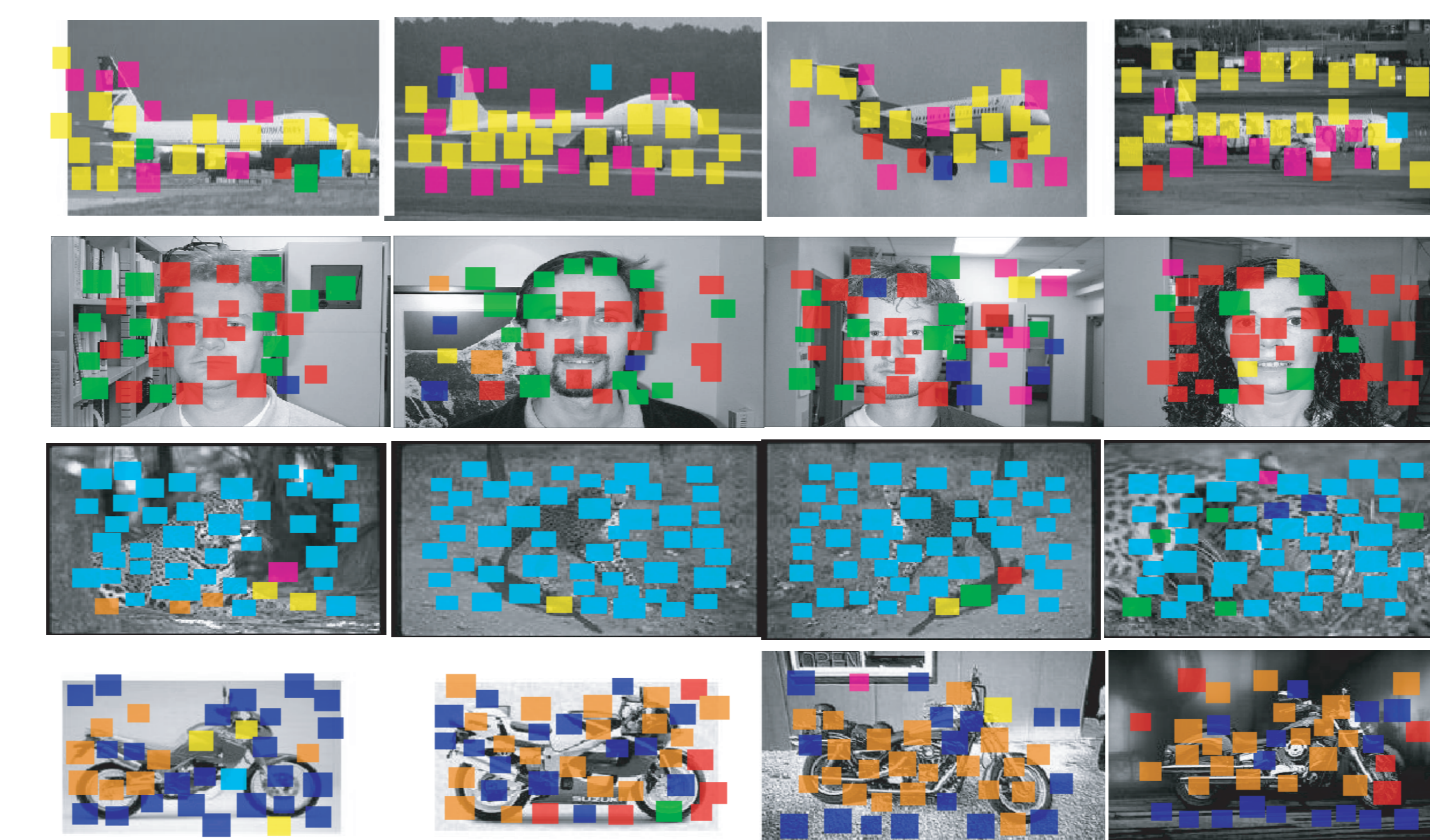


## 7. Experiment 1: Caltech 4 dataset

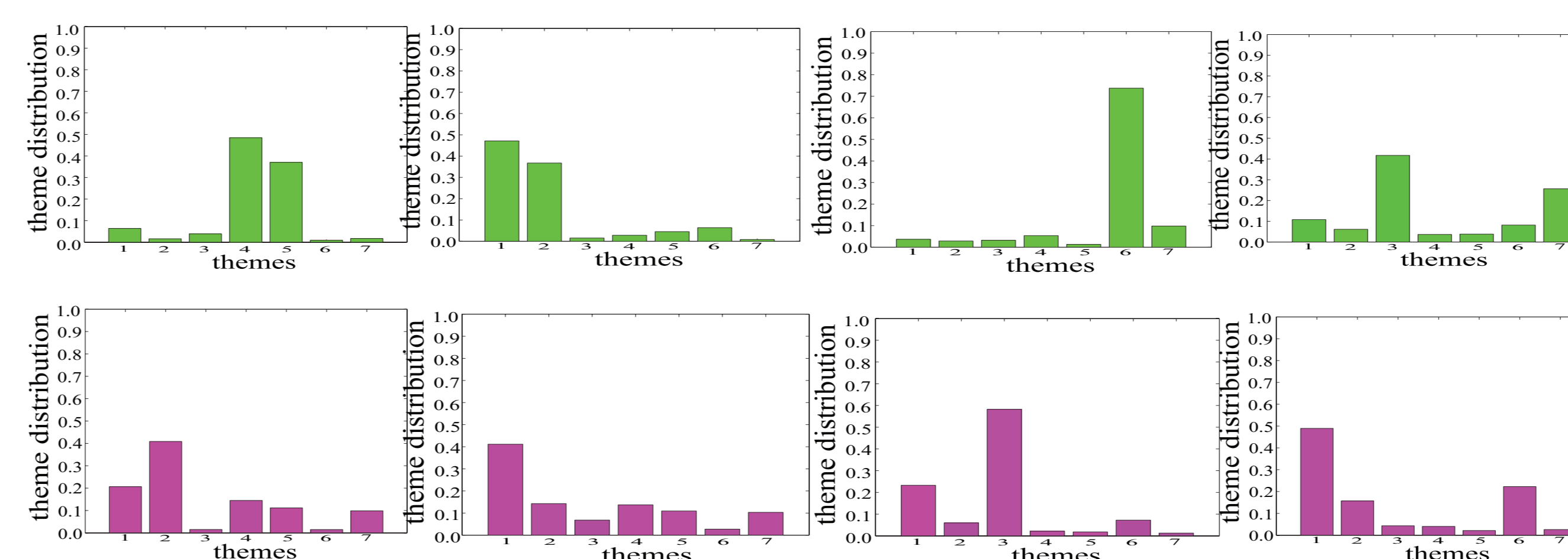
The five most dependent patches for each category.



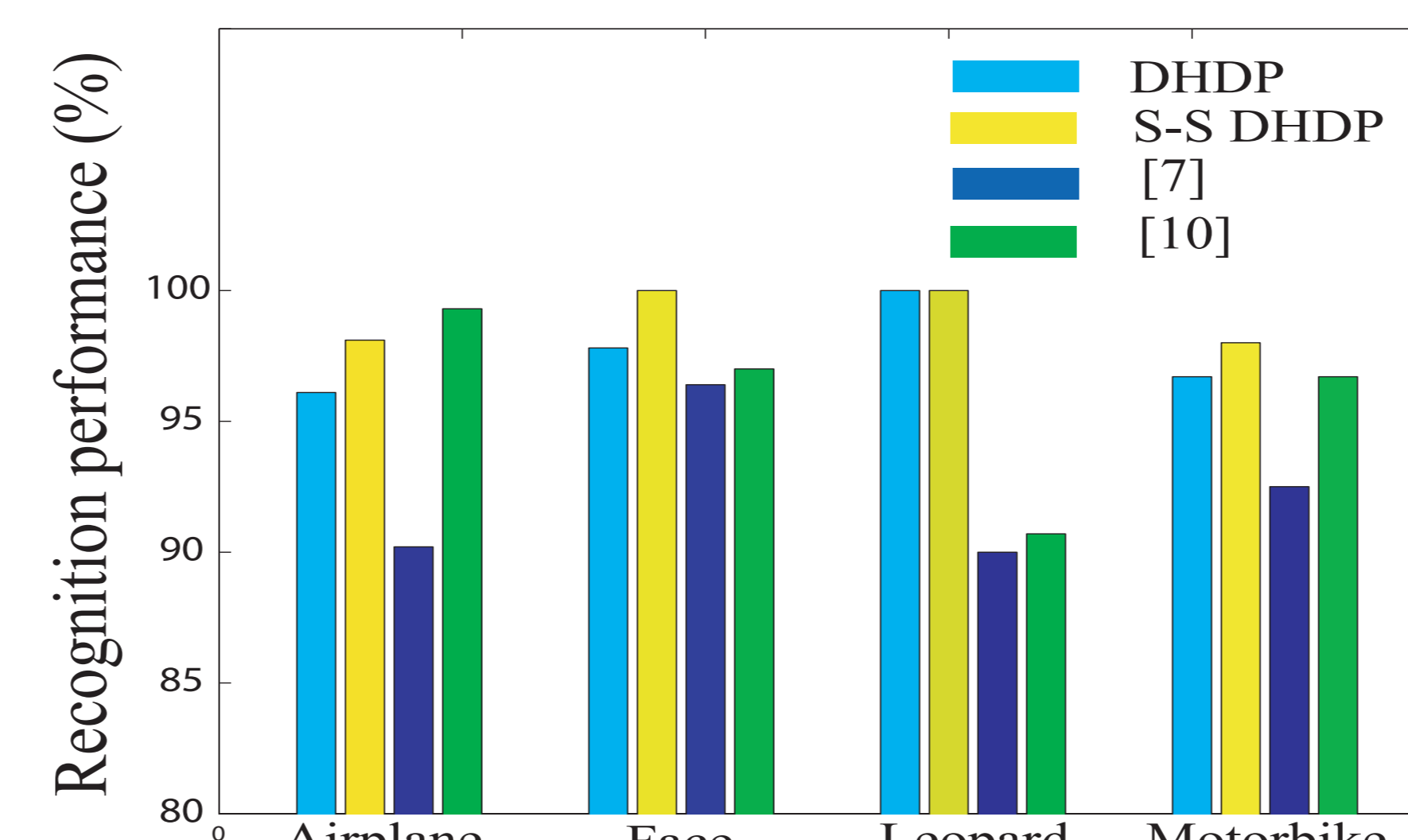
Local patches from different themes.



Theme distribution of DHDP and HDP.



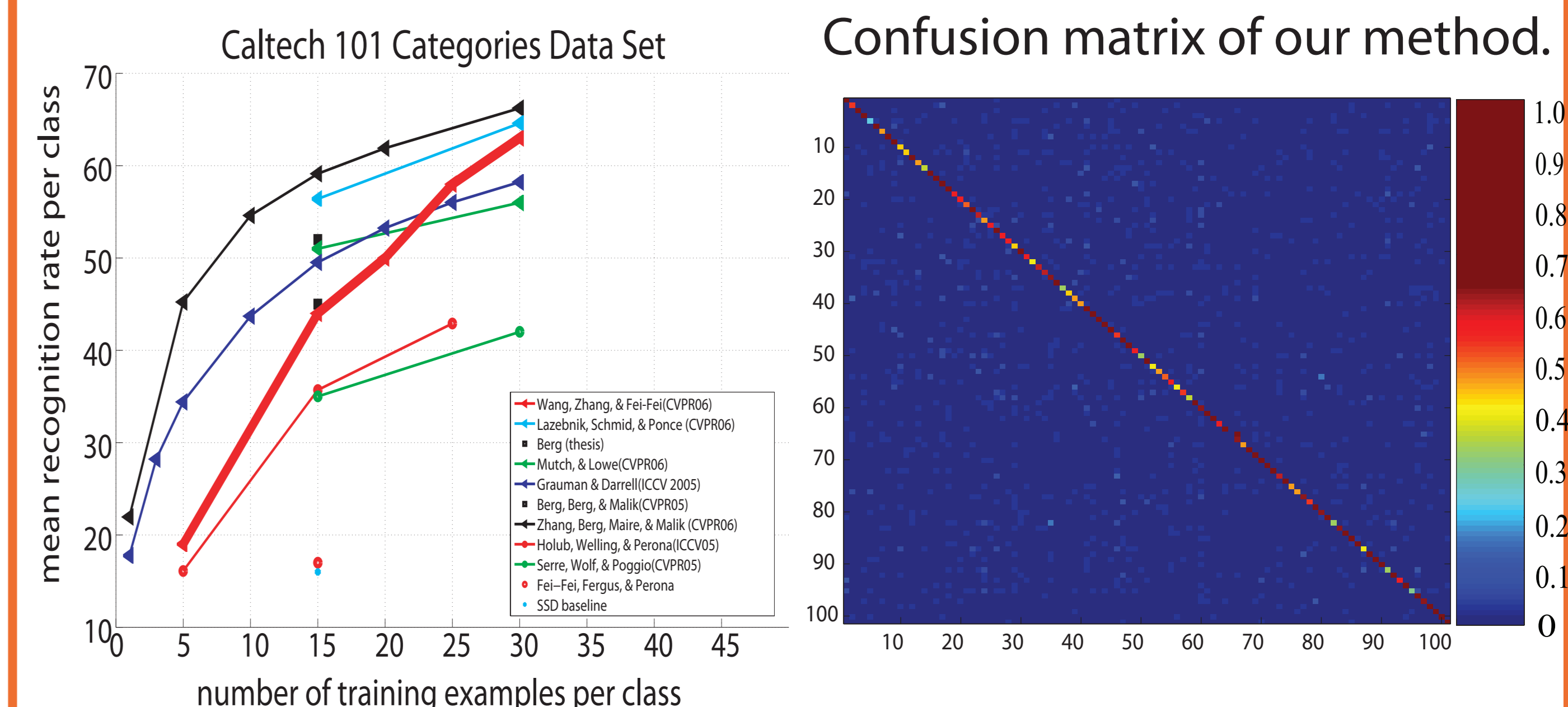
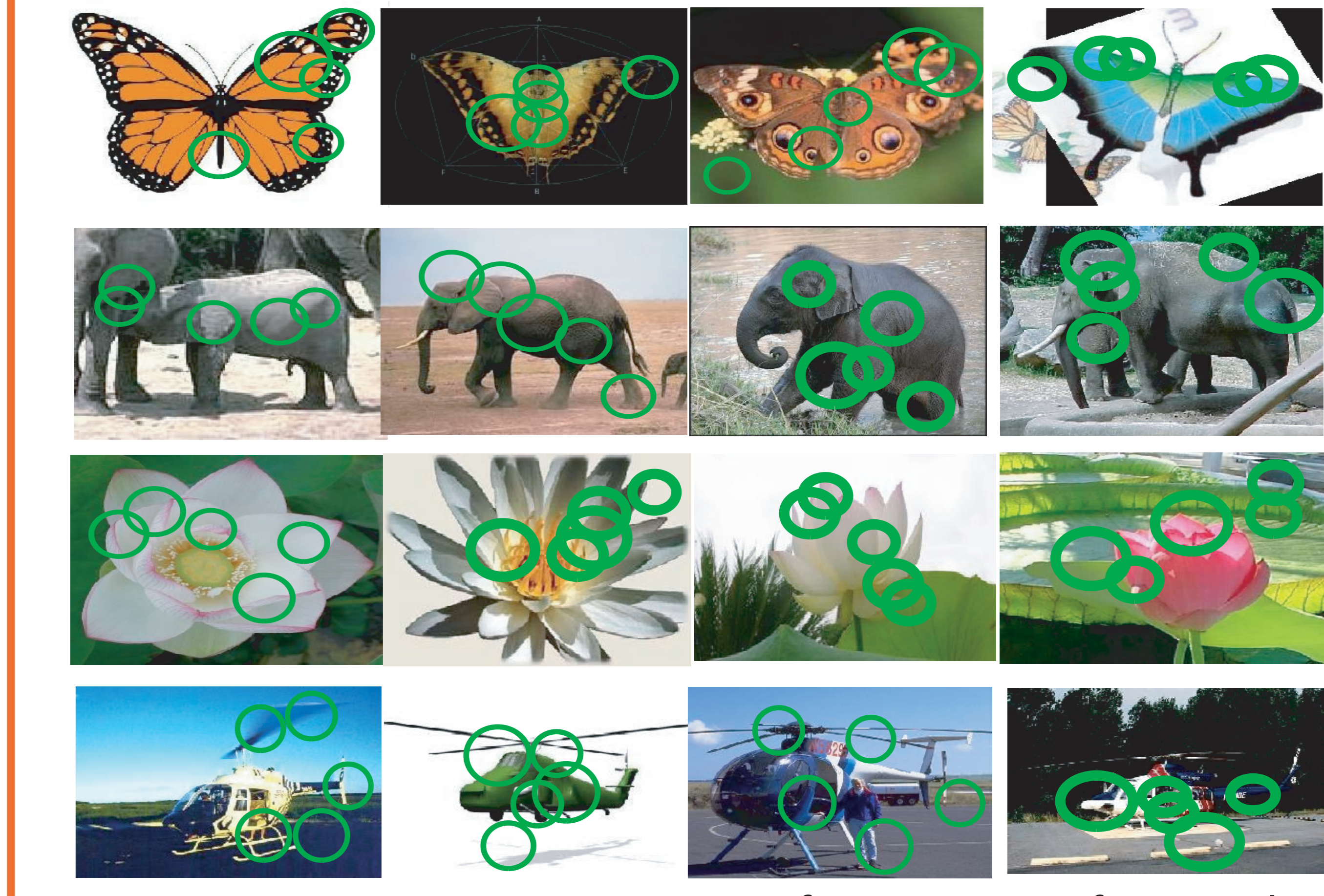
Performance comparison of DHDP, semi-supervised DHDP to [7] and [10].



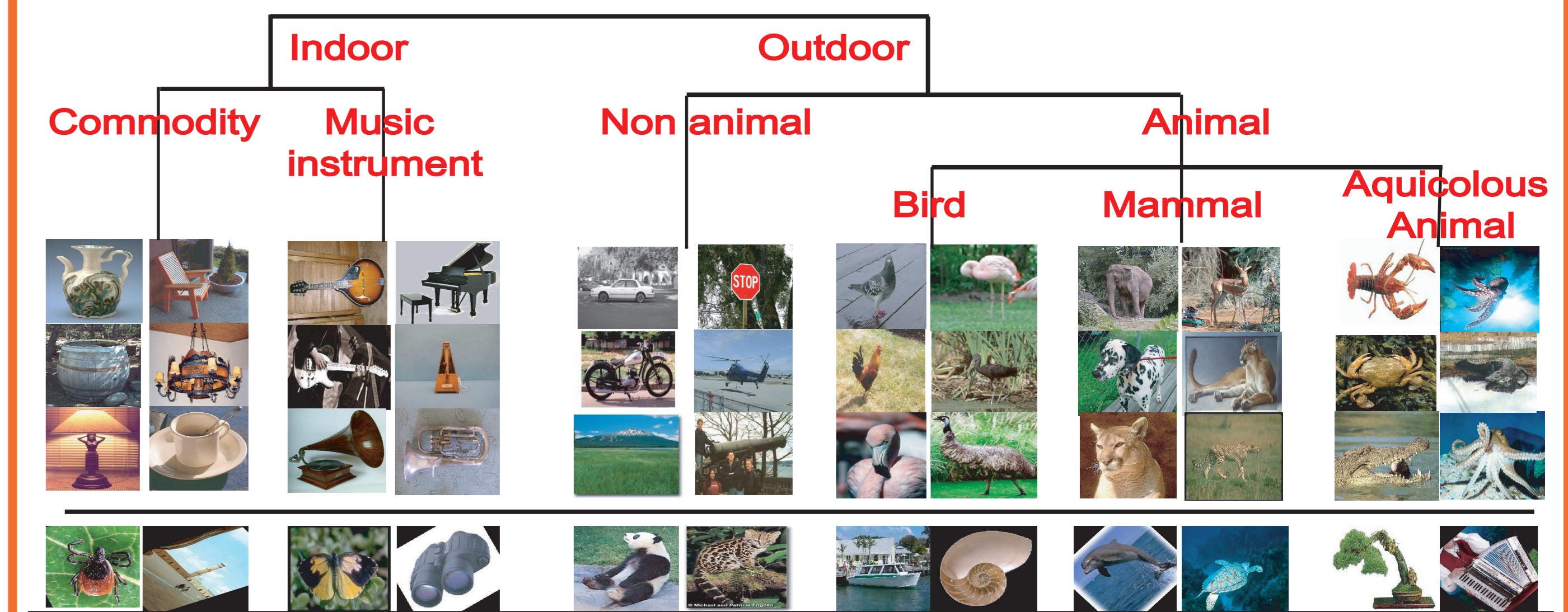
Reference:  
 [7] R. Fergus, P. Perona, and A. Zisserman. Object class recognition by unsupervised scale-invariant learning. IJCV (submitted), 2005.  
 [10] A. Holub, M. Welling, and P. Perona. Combining generative models and fisher kernels for object class recognition. In ICCV, 2005.

## 8. Experiment 2: Caltech 101 dataset

The five most dependent patches for each category.



Extracted taxonomy on the Caltech 101 categories data set.



**Commodity:** Wheelchair Chair Chandelier Windsor\_chair Ceiling\_fan Lamp Umbrella Scissors Watch Headphone Wrench Cup Ewer Barrel Inline\_skate Cellphone Pizza Garfield Starfish Revolver Tick Triobite Minaret Menorah Saxophone Anchor Pyramid Scorpion Brain

**Music instrument:** Metronome Mandolin Grand\_piano Euphonium Electric\_guitar Laptop Camera Gramophone Butterfly Binocular Stapler

**Non animal:** Car-side Helicopter Airplanes Motorbike Cannon Schooner Pagoda Background Stegosaurus Lotus Sunflower Strawberry Joshua\_tree Water\_lilly Yin\_yang Dollar\_Bill Soccer\_ball Stop\_sign Buddha Faces Sea\_horse Okapi Panda Hedgehog Beaver Lobster Wild\_cat

**Bird:** Rooster Platypus Flamingo Flamingo\_head Nautilus Pigeon Ibis Emu Rhino llama Ferry Kangaroo

**Mammal:** Elephant Cougar\_body Gerenuk Cougar\_face Dalmatian Leopard Dolphin Snoopy Hawksbill

**Aquicolous Animal:** Octopus Crocodile\_head Crab Crocodile Crayfish Bonsai Brontosaurus Accordion

Main reference:  
 [1] G. Wang, Y. Zhang, and L. Fei-Fei. Using dependent regions for object categorization in a generative framework. IEEE Comp. Vis. Patt. Recog. 2006.  
 [2] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei. Hierarchical Dirichlet processes. Journal of the American Statistical Association, 2006.

Corresponding author: feifeili@uiuc.edu