Fine-Grained Image Analysis

ICME Tutorial

Xiu-Shen WEI

Megvii Research Nanjing, Megvii Inc.

Jul. 8, 2019 Shanghai, China

Outline

- ☑ Background about CV, DL, and image analysis
- ☑ Introduction of fine-grained image analysis
- ☑ Fine-grained image retrieval

Part 1

- ☑ Fine-grained image recognition
- ☑ Other computer vision tasks related to fine-grained image analysis
- ☑ New developments of fine-grained image analysis

Part 2

Part I

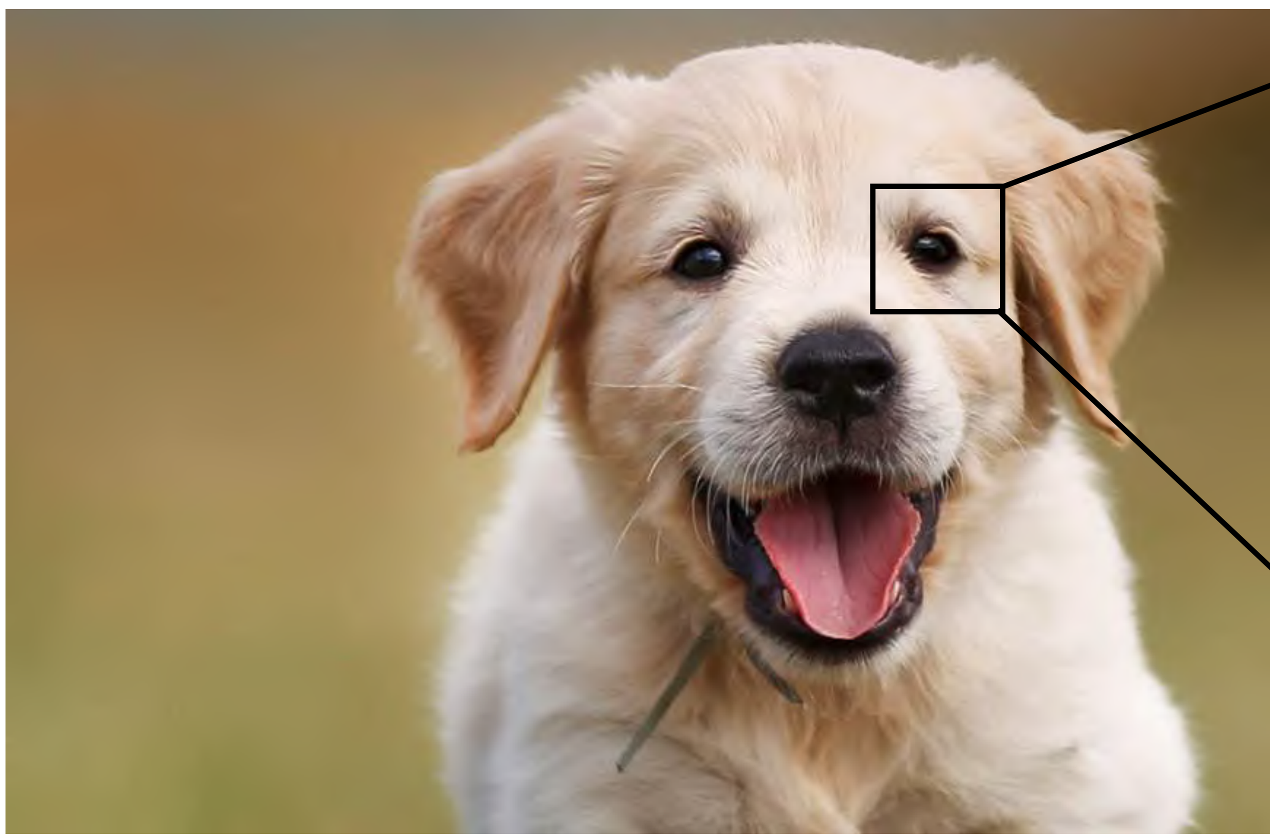
- Background
  - A brief introduction of computer vision
  - Traditional image recognition and retrieval
  - Deep learning and convolutional neural networks

- Introduction
  - Fine-grained images *vs.* generic images
  - Various real-world applications of fine-grained images
  - Challenges of fine-grained image analysis
  - Fine-grained benchmark datasets

- Fine-grained image retrieval
  - Fine-grained image retrieval based on hand-crafted features
  - Fine-grained image retrieval based on deep learning

Background

What is computer vision?



What we see

0	3	2	5	4	7	6	9	8
3	0	1	2	3	4	5	6	7
2	1	0	3	2	5	4	7	6
5	2	3	0	1	2	3	4	5
4	3	2	1	0	3	2	5	4
7	4	5	2	3	0	1	2	3
6	5	4	3	2	1	0	3	2
9	6	7	4	5	2	3	0	1
8	7	6	5	4	3	2	1	0

What a computer sees

Background (con't)

Why study computer vision?

- CV is useful
- CV is interesting
- CV is difficult
- …

Crowds and occlusions

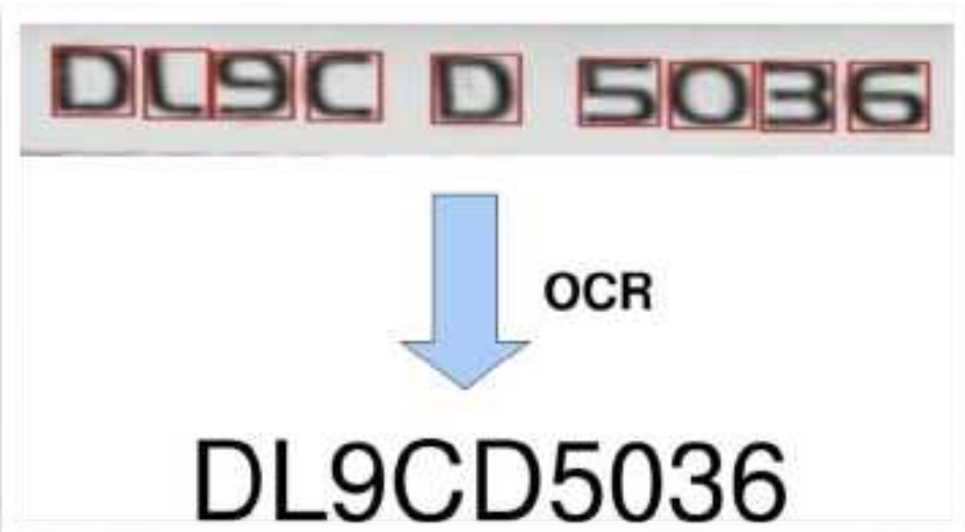
"man in black shirt is playing guitar."

Image captioning

Finger reader

Background (con't)

Successes of computer vision to date



Optical character recognition

Biometric systems

Face recognition

Self-driving cars

Background (con't)

Top-tier CV conferences/journals and prizes

2018 CVPR SALT LAKE CITY • JUNE 18-22

ICCV 17

ECCV 2018

Marr Prize

Background (con't)

Traditional image recognition and image retrieval

Data → Feature extraction → Learning algorithm

Recognition:
SVM
Random Forest
…

Retrieval:
k-NN
…

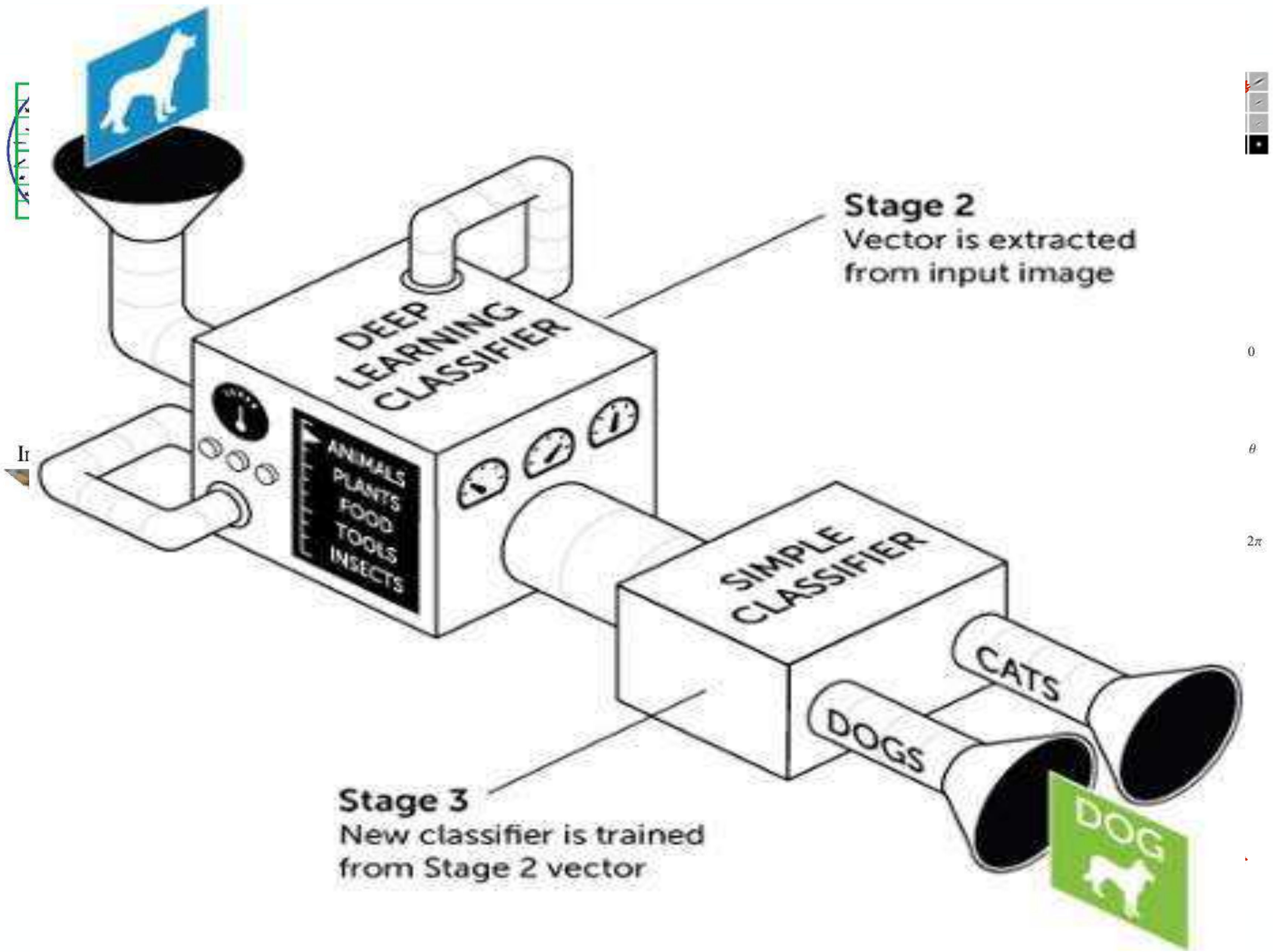
Image

vision features

Recognition

Background (con't)

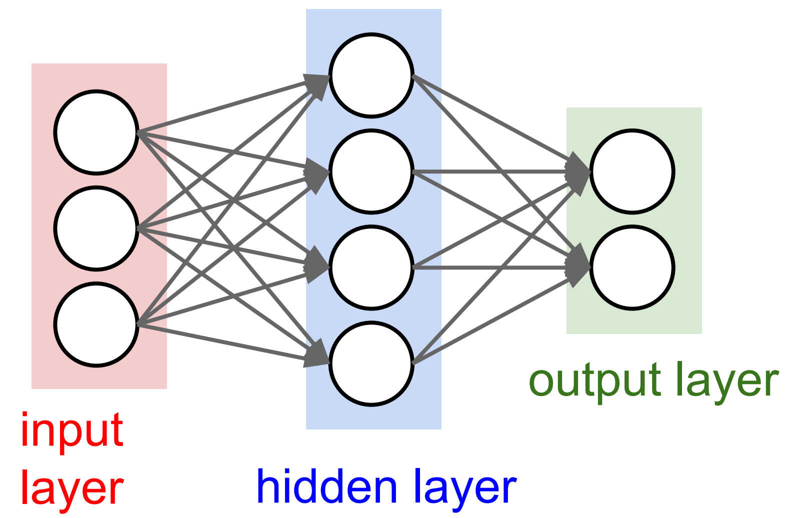
Computer vision features



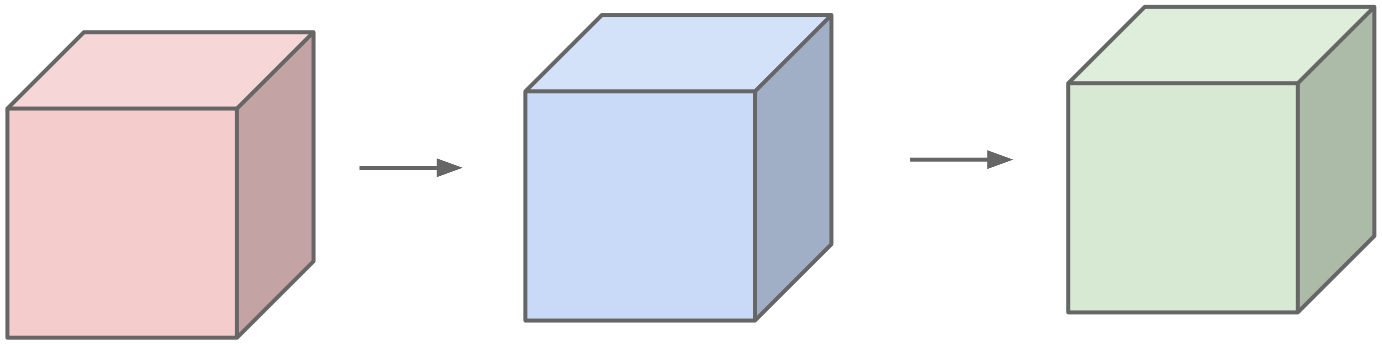
Background (con't)

Deep learning and convolutional neural networks

Before:



Now:

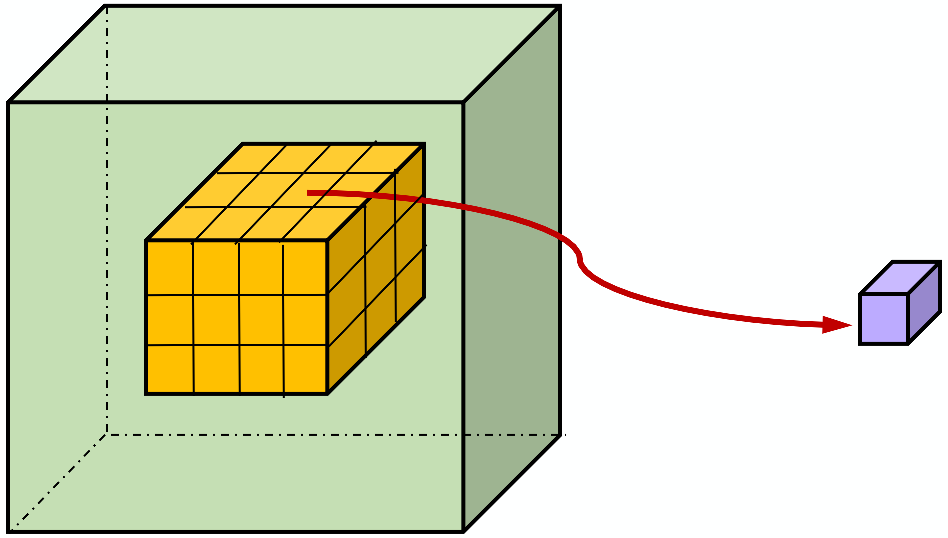


"Rome was not built in one day!"

Background (con't)

Deep learning and convolutional neural networks

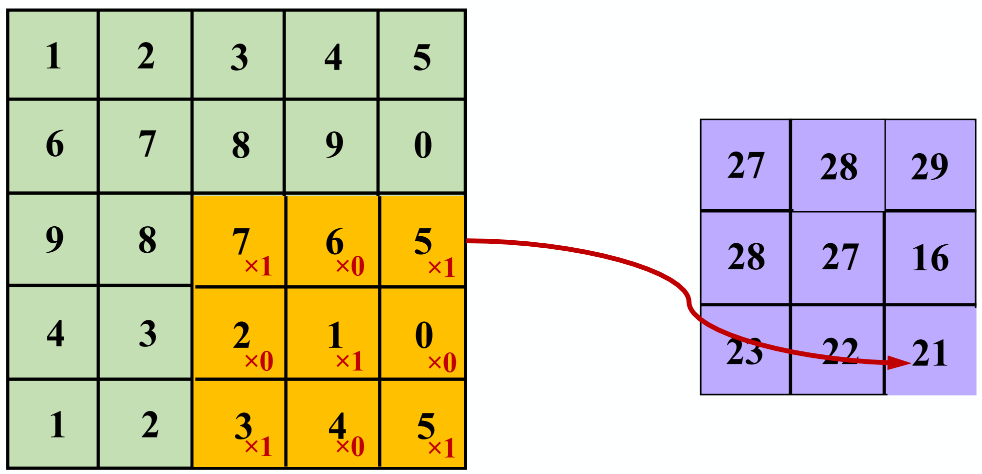
All neural net activations arranged in 3-dimension:



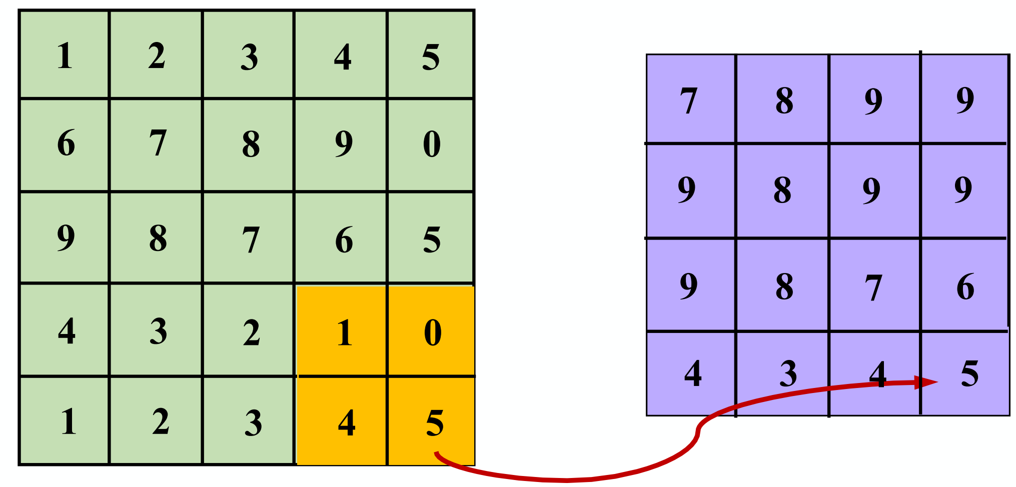
$$y_{i^{l+1},j^{l+1},d} = \sum_{i=0}^{H} \sum_{j=0}^{W} \sum_{d^l=0}^{D^l} f_{i,j,d^l,d} \times x^l_{i^{l+1}+i,j^{l+1}+j,d^l}$$

Background (con't)

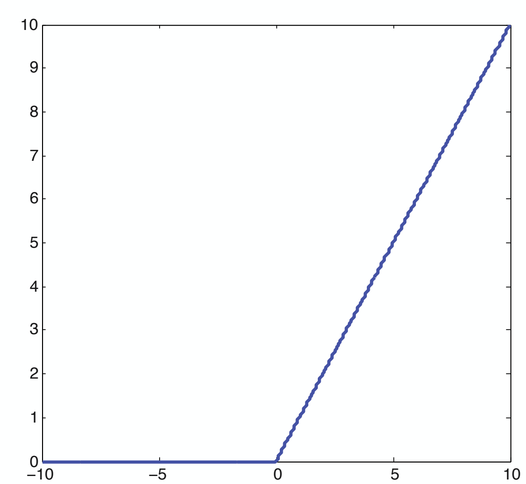
Unit processing of CNNs



Convolution



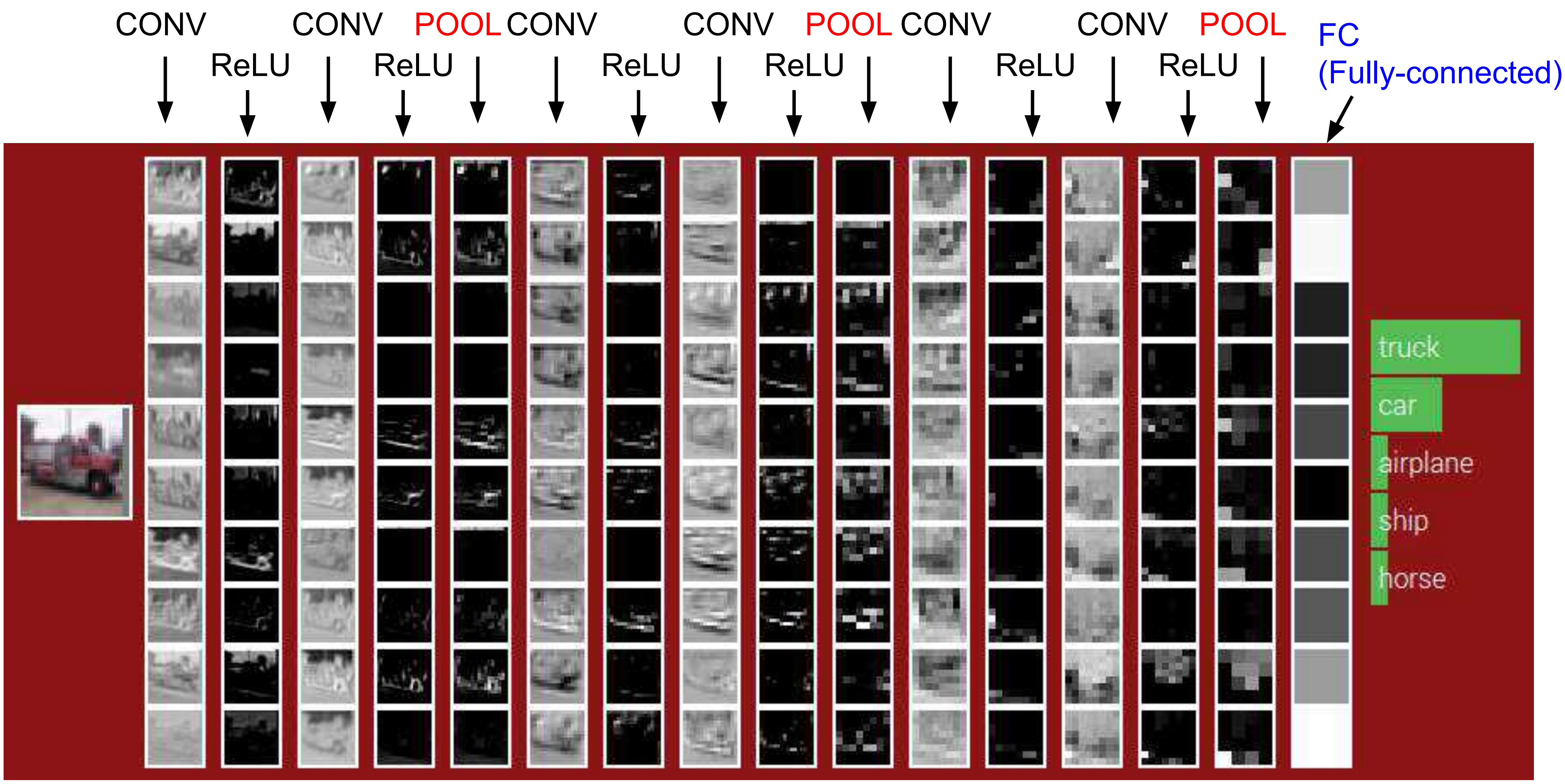
Pooling



Non-linear activation function

Background (con't)

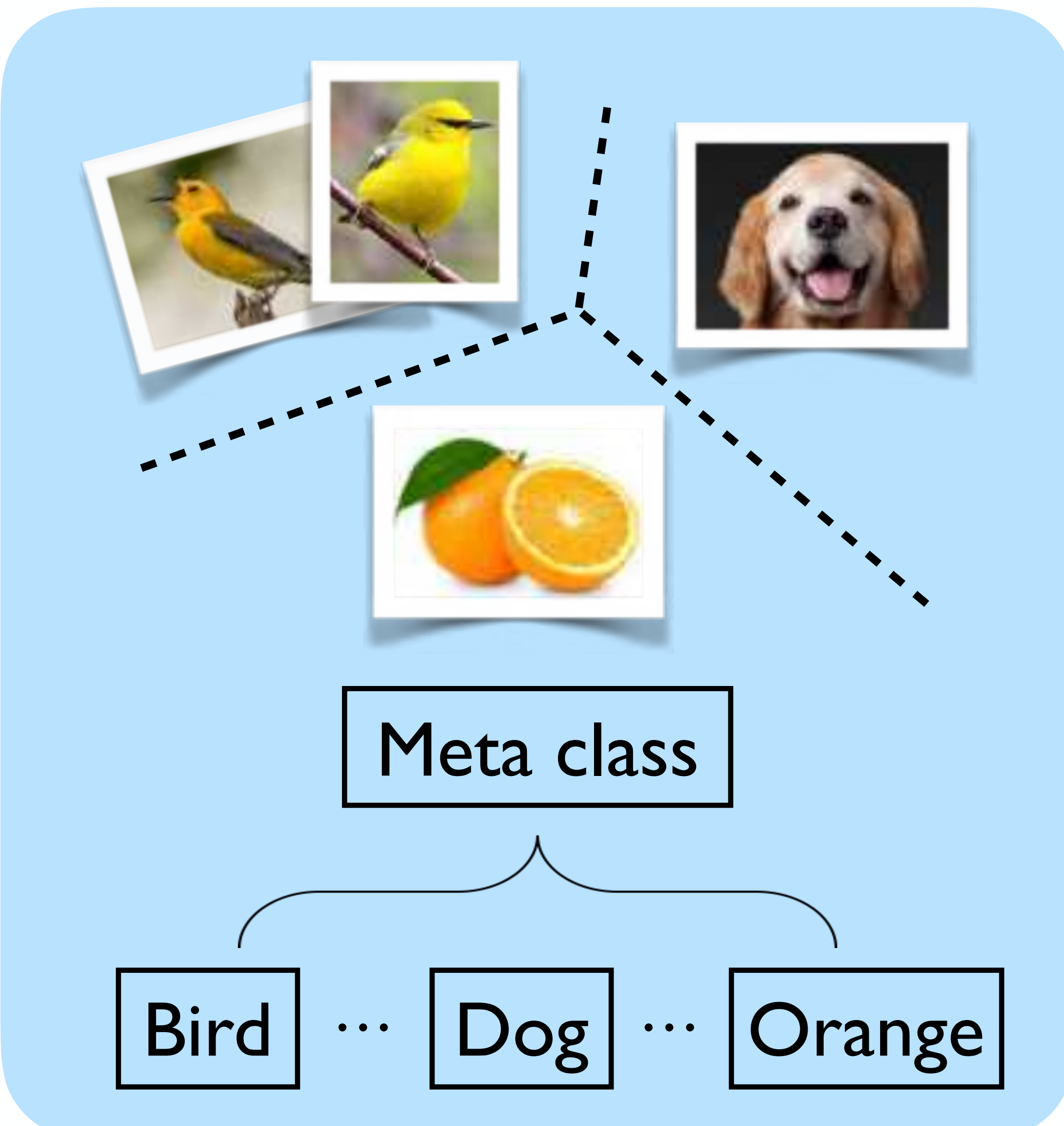
CNN architecture



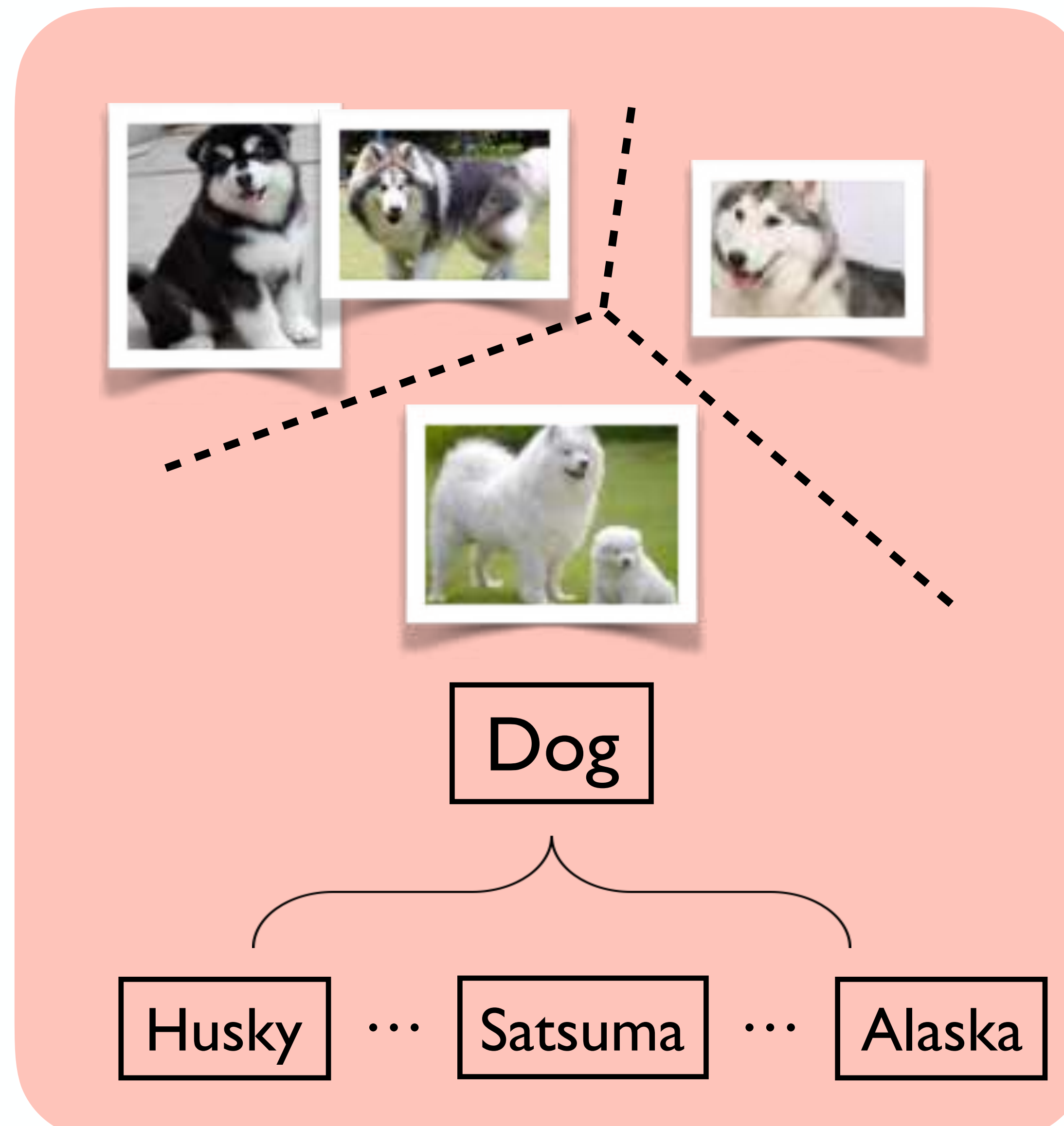
Figures are courtesy of L. Fei-Fei.

Introduction

Fine-grained images *vs.* generic images



Traditional image recognition
(Coarse-grained)



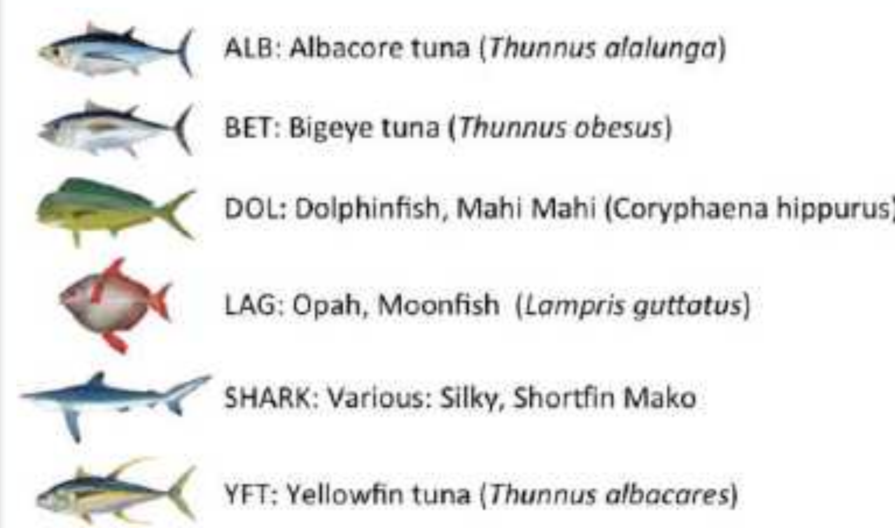
Fine-grained image recognition

Introduction (con't)

Various real-world applications

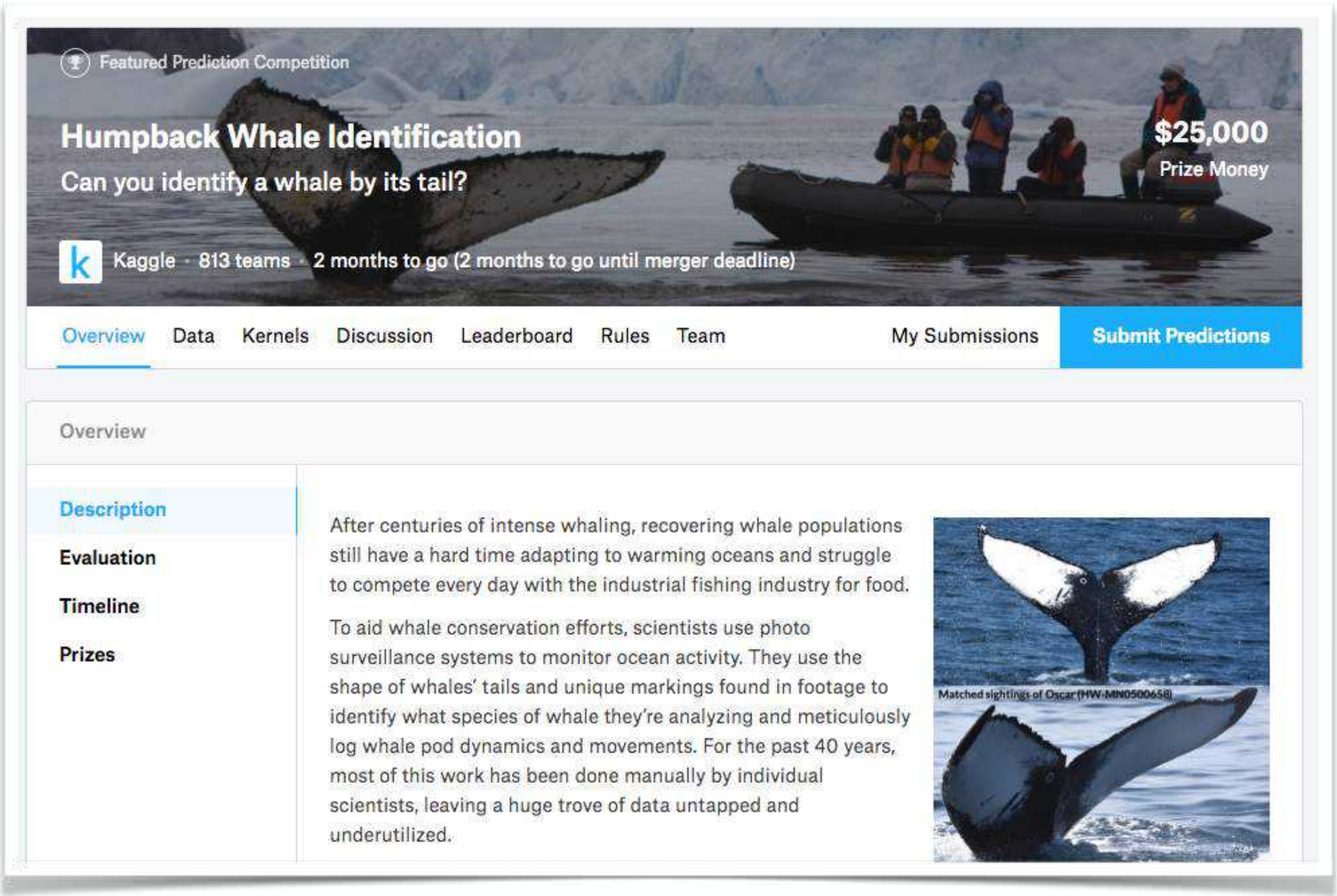
Can you detect and classify species of fish?

Nearly half of the world depends on seafood for their main source of protein. In the Western and Central Pacific, where 60% of the world's tuna is caught, illegal, unreported, and unregulated fishing practices are threatening marine ecosystems, global seafood supplies and local livelihoods. The Nature Conservancy is working with local, regional and global partners to preserve this fishery for the future.



Introduction (con't)

Various real-world applications



Introduction (con't)

Various real-world applications

Results

1. Megvii Research Nanjing
   a. Error = 0.10267
2. Alibaba Machine Intelligence Technology Lab
   a. Error = 0.11315
3. General Dynamics Mission Systems
   a. Error = 0.12678


FGVC6 iNat2019

This certificate is awarded to

Bo-Yan Zhou, Bo-Rui Zhao, Quan Cui, Yan-Ping Xie, Zhao-Min Chen, Ren-Jie Song, and Xiu-Shen Wei
Megvii Research Nanjing

winners of the iNaturalist 2019 image classification challenge held in conjunction with the FGVC workshop at CVPR 2019.

Sponsored by Microsoft

Introduction (con't)

Various real-world applications

Herbarium Challenge 2019

Kiat Chuan Tan[1], Yulong Liu[1], Barbara Ambrose[2], Melissa Tulig[2], Serge Belongie[1,3]

[1]Google Research, [2]New York Botanical Garden, [3]Cornell Tech

Google Research

CORNELL TECH

Herbarium Challenge 2019 (top

- #1 Megvii Research Nanjing (89.8%)
  - Boyan Zhou, Quan Cui, Borui Zhao, Yanping Xie, Renjie Song, **Xiu-Shen Wei**
- #2 PEAK (89.1%)
  - Chunqiao Xu, Shao Zeng, Qiule Sun, Shuyu Ge, Peihua Li (Dalian University of Technology)
- #3 Miroslav Valan (89.0%)
  - Swedish Museum of Natural History
- #4 Hugo Touvron (88.9%)
  - Hugo Touvron and Andrea Vedaldi (Facebook AI Research)

Introduction (con't)

Various real-world applications

Introduction (con't)

Various real-world applications

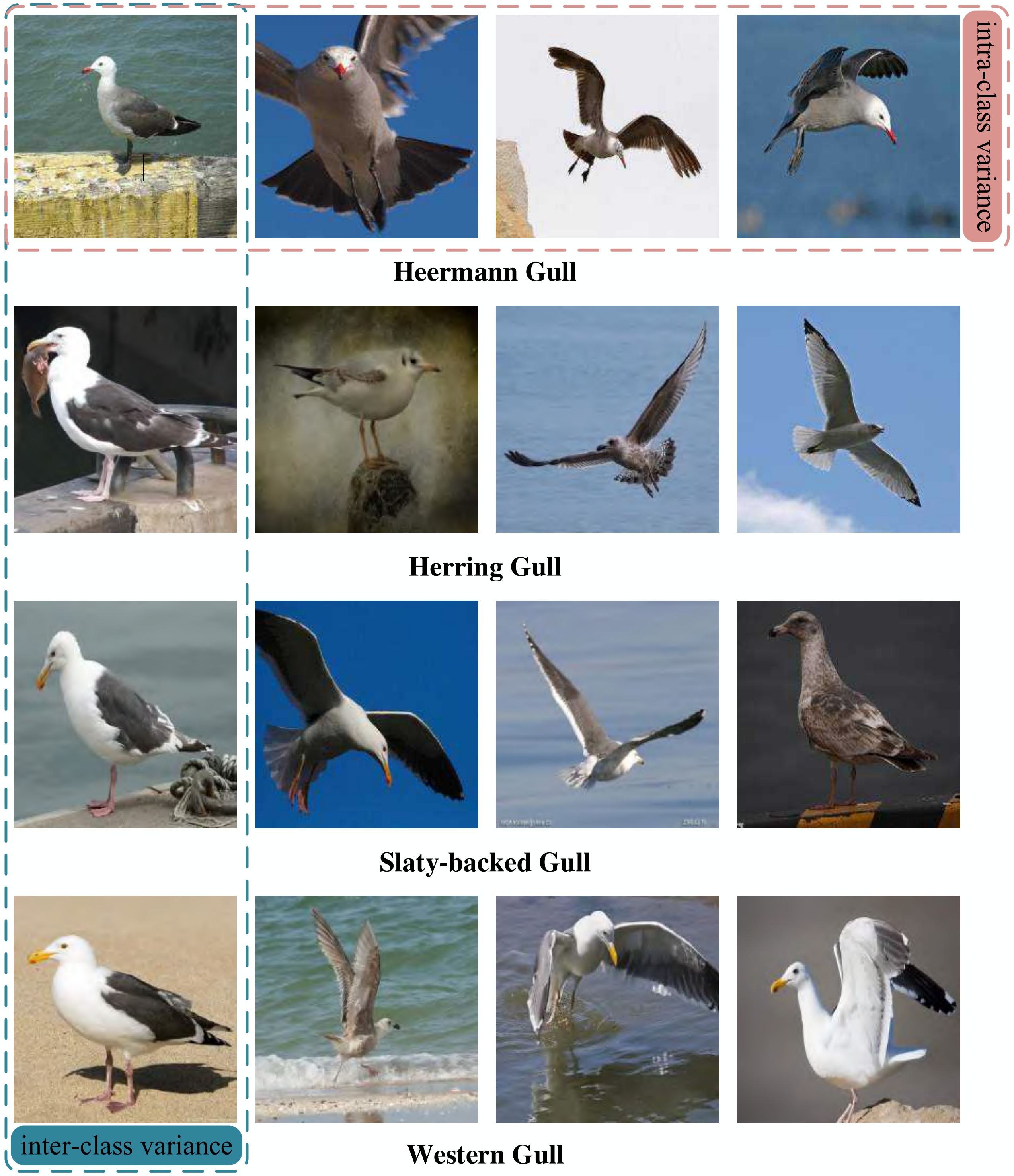
Introduction (con’t)

Various real-world applications

Introduction (con’t)

Challenge of fine-grained image analysis

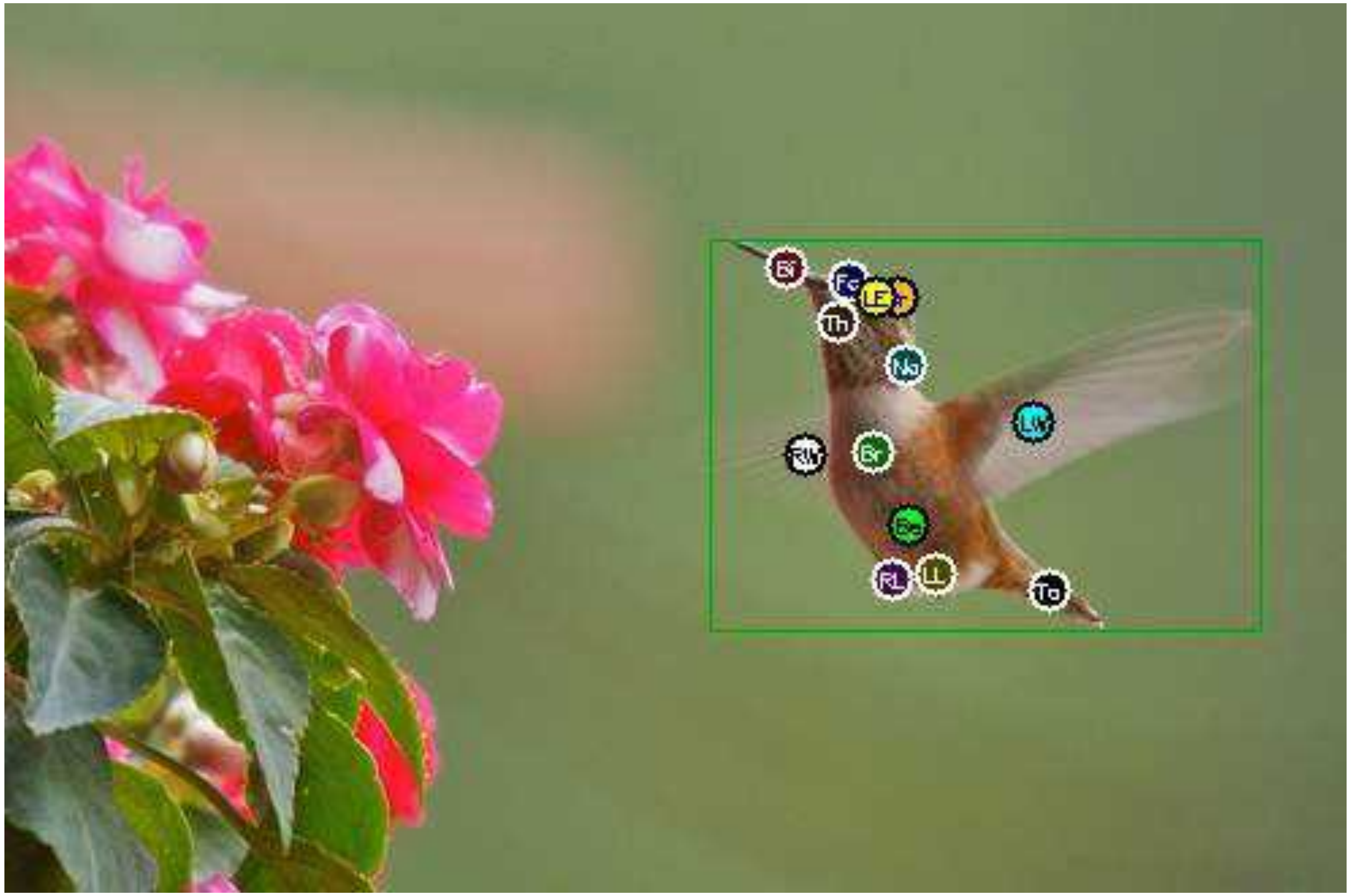
Small inter-class variance
Large intra-class variance



Figures are courtesy of [X. He and Y. Peng, CVPR 2017].

Introduction (con't)

The key of fine-grained image analysis



Introduction (con't)

Fine-grained benchmark datasets

CUB200-2011

- 11,788 images, 200 fine-grained classes

Introduction (con't)

Fine-grained benchmark datasets

Stanford Dogs

- 20,580 images
- 120 fine-grained classes

Introduction (con't)

Fine-grained benchmark datasets

Oxford Flowers

- 8,189 images, 102 fine-grained classes

[M.-E. Nilsback and A. Zisserman, CVGIP 2008]

Introduction (con't)

Fine-grained benchmark datasets

Aircrafts

- 10,200 images
- 100 fine-grained classes

Introduction (con't)

Fine-grained benchmark datasets

Stanford Cars

- 16,185 images, 196 fine-grained classes

[J. Krause et al., ICCV Workshop 2013]

Introduction (con't)

Fine-grained image analysis is hot …

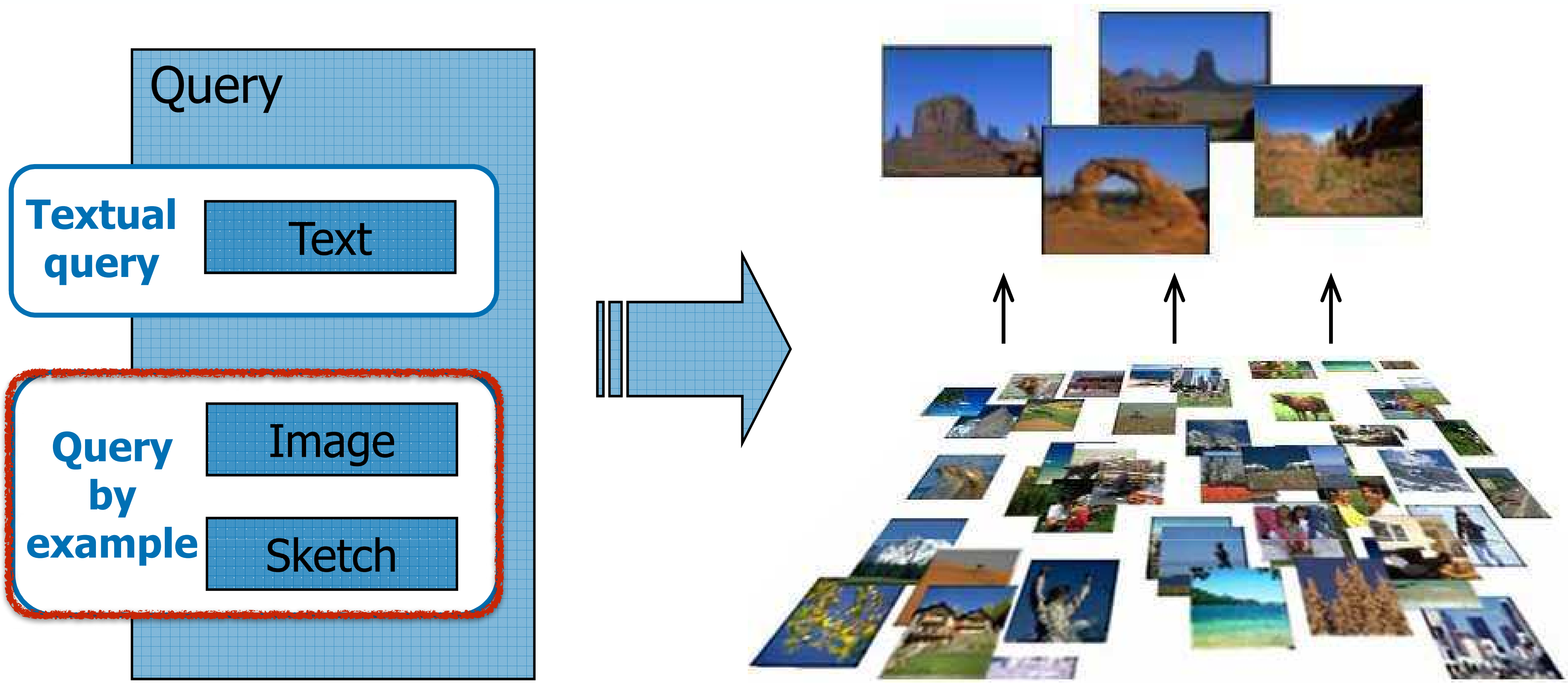
- Many papers published on top-tier conf./journals
  - CVPR, ICCV, ECCV, IJCAI, etc.
  - TPAMI, IJCV, TIP, etc.

- Many frequently held workshops
  - Workshop on Fine-Grained Visual Categorization
  - …

- Many academic challenges about fine-grained tasks
  - The Nature Conservancy Fisheries Monitoring
  - iFood Classification Challenge
  - iNature Classification Challenge
  - …

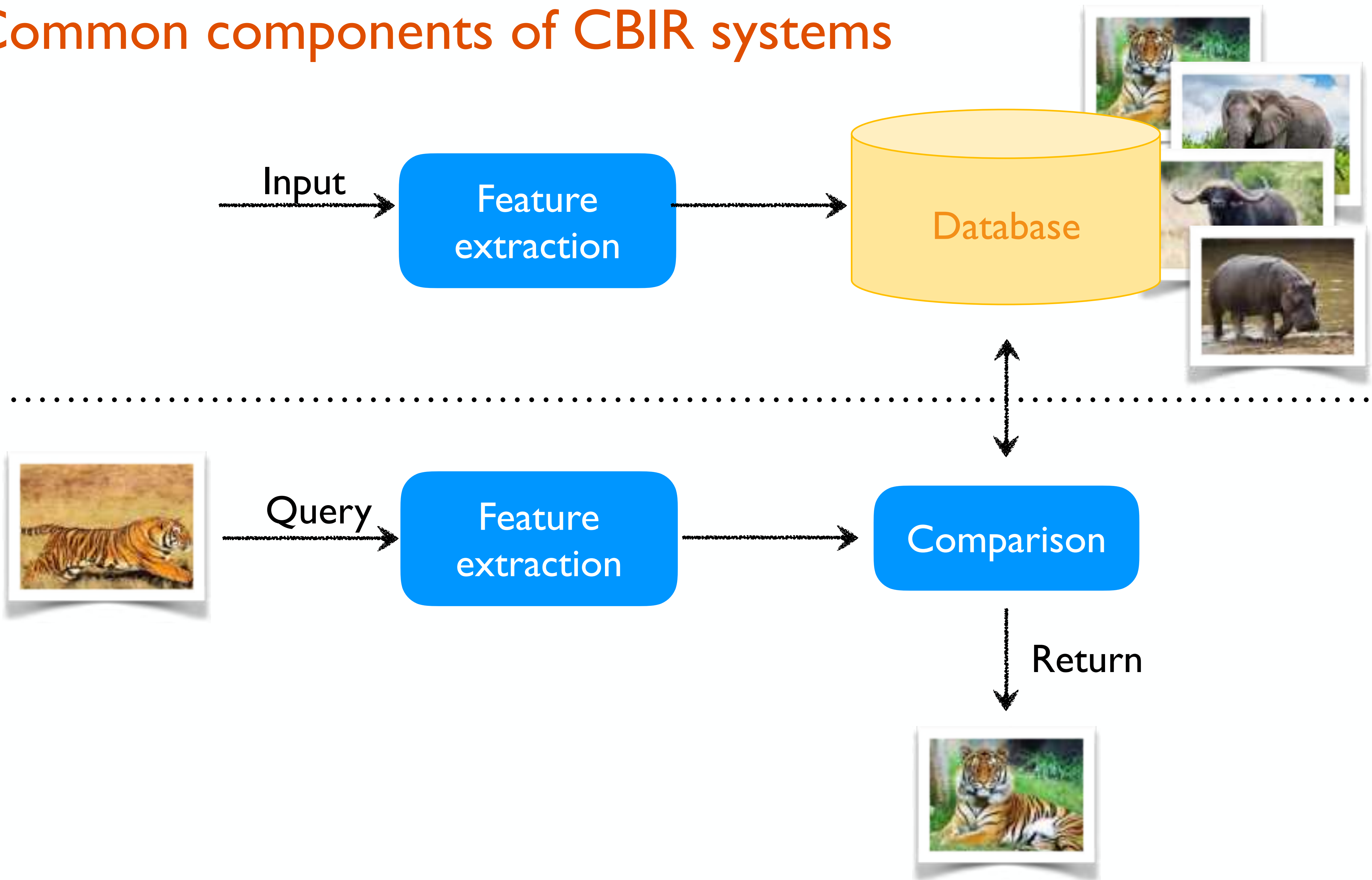
Fine-grained image retrieval

Image Retrieval (IR)



Fine-grained image retrieval (con't)

Common components of CBIR systems



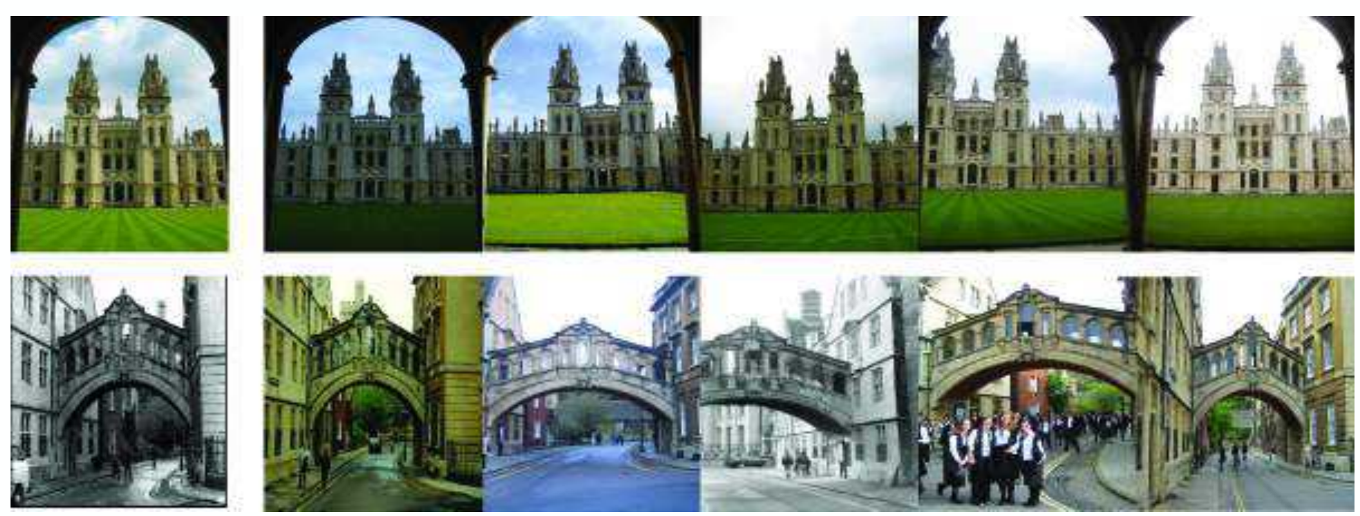
Fine-grained image retrieval (con't)

Deep learning for image retrieval

Fine-grained image retrieval (con't)

FGIR vs. General-purposed IR

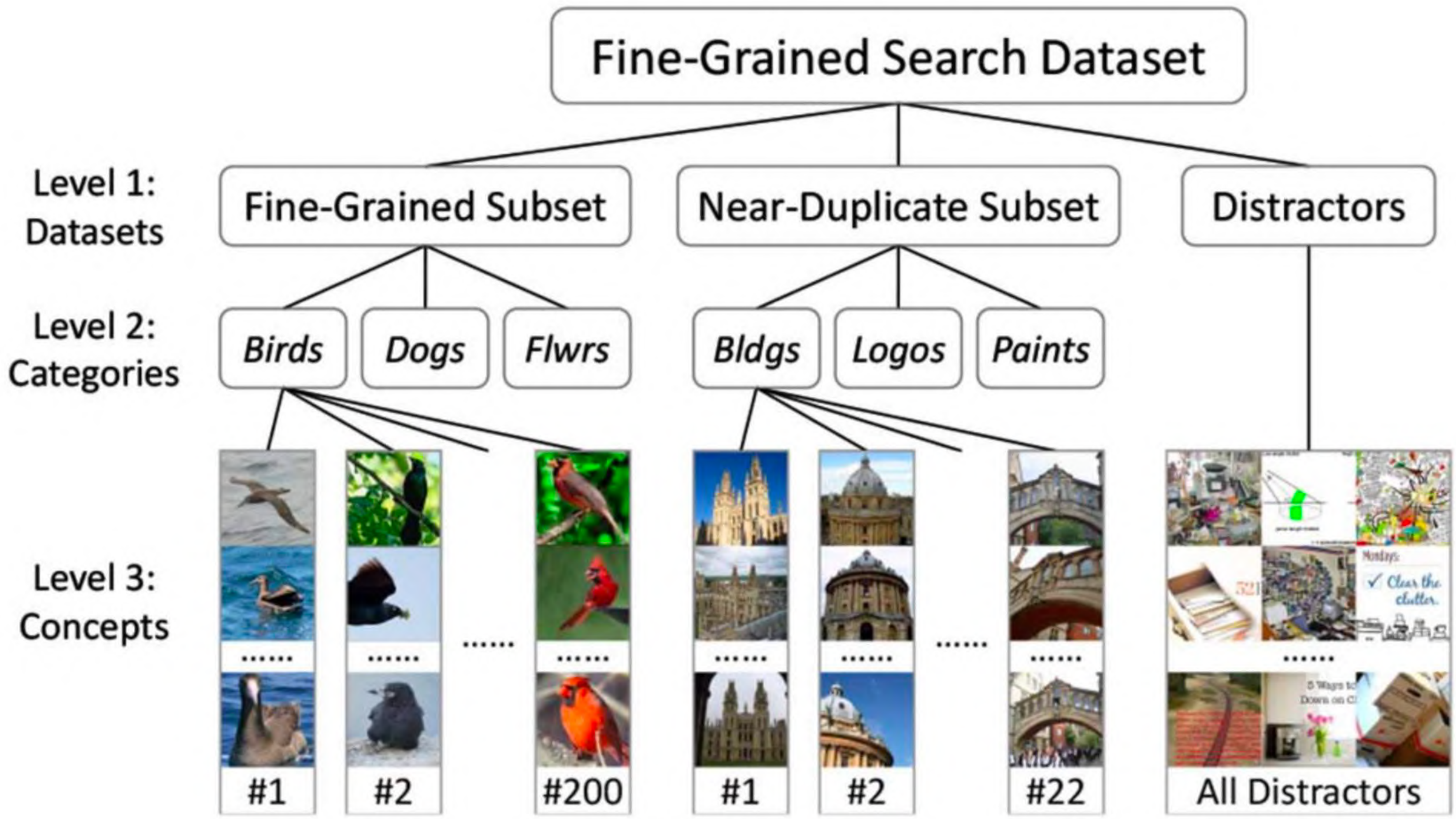
(a) Fine-grained image retrieval. Two examples ("Mallard" and "Rolls-Royce Phantom Sedan 2012") from the *CUB200-2011* [10] and *Cars* [11] datasets, respectively.



(b) General image retrieval. Two examples from the *Oxford Building* [12] dataset.

Fine-grained image retrieval (con't)

FGIR based on hand-crafted features



[Xie et al., IEEE TMM 2015]

Fine-grained image retrieval (con't)

Selective Convolutional Descriptor Aggregation (SCDA)

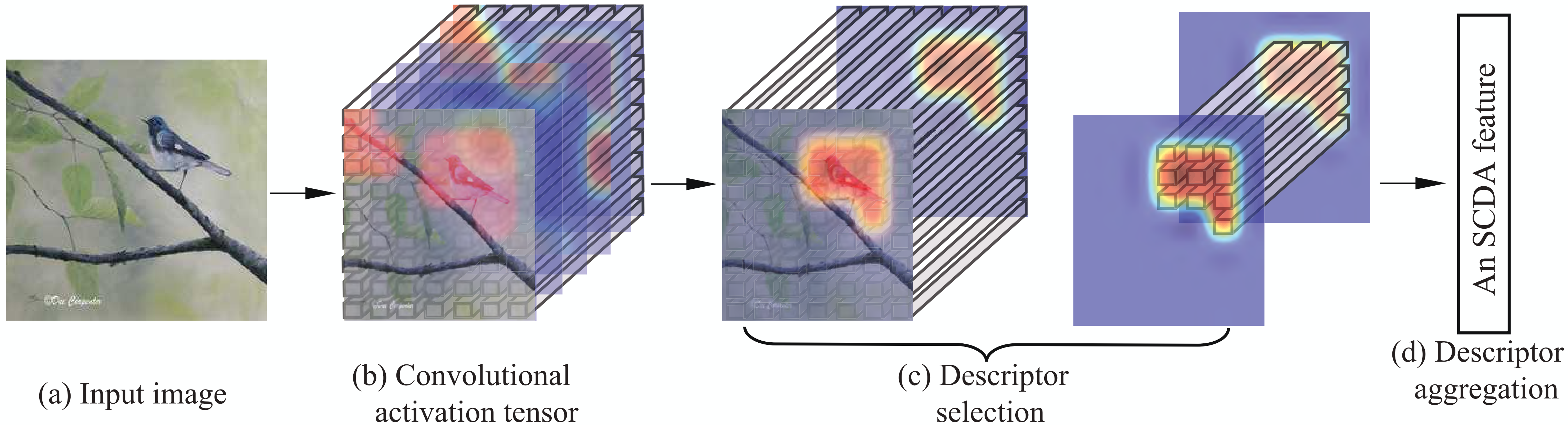
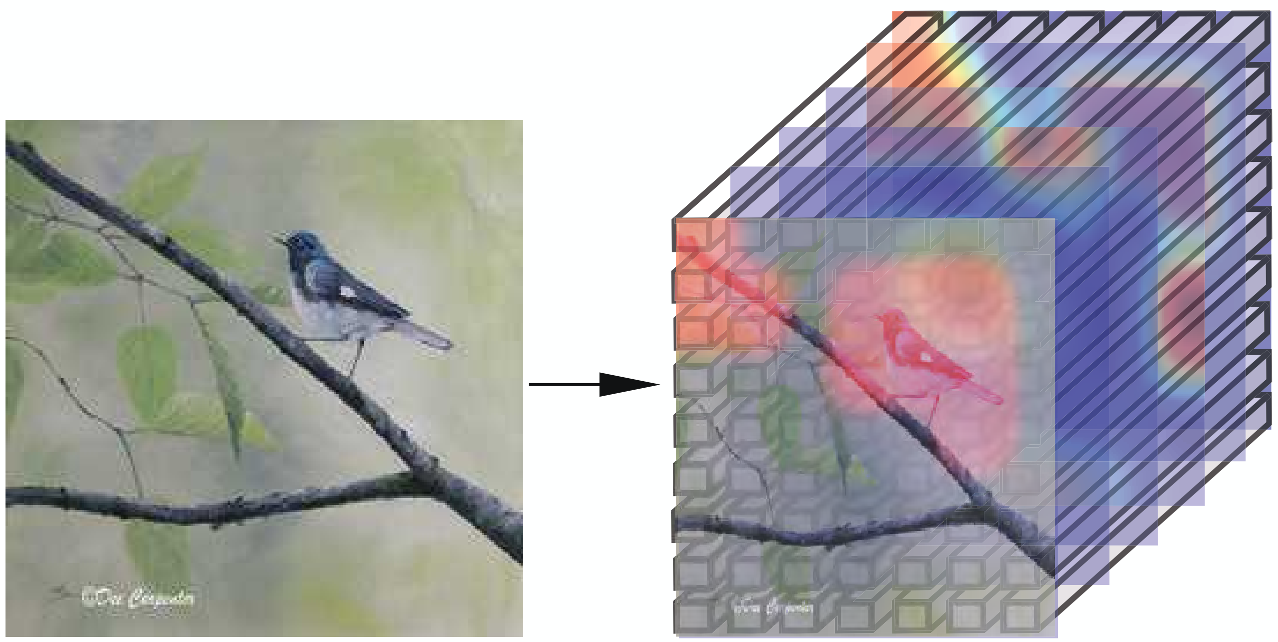


Figure 1. Pipeline of the proposed SCDA method. (Best viewed in color.)

Fine-grained image retrieval (con't)

Notations



(a) Input image

(b) Convolutional activation tensor

$h \times w \times d$

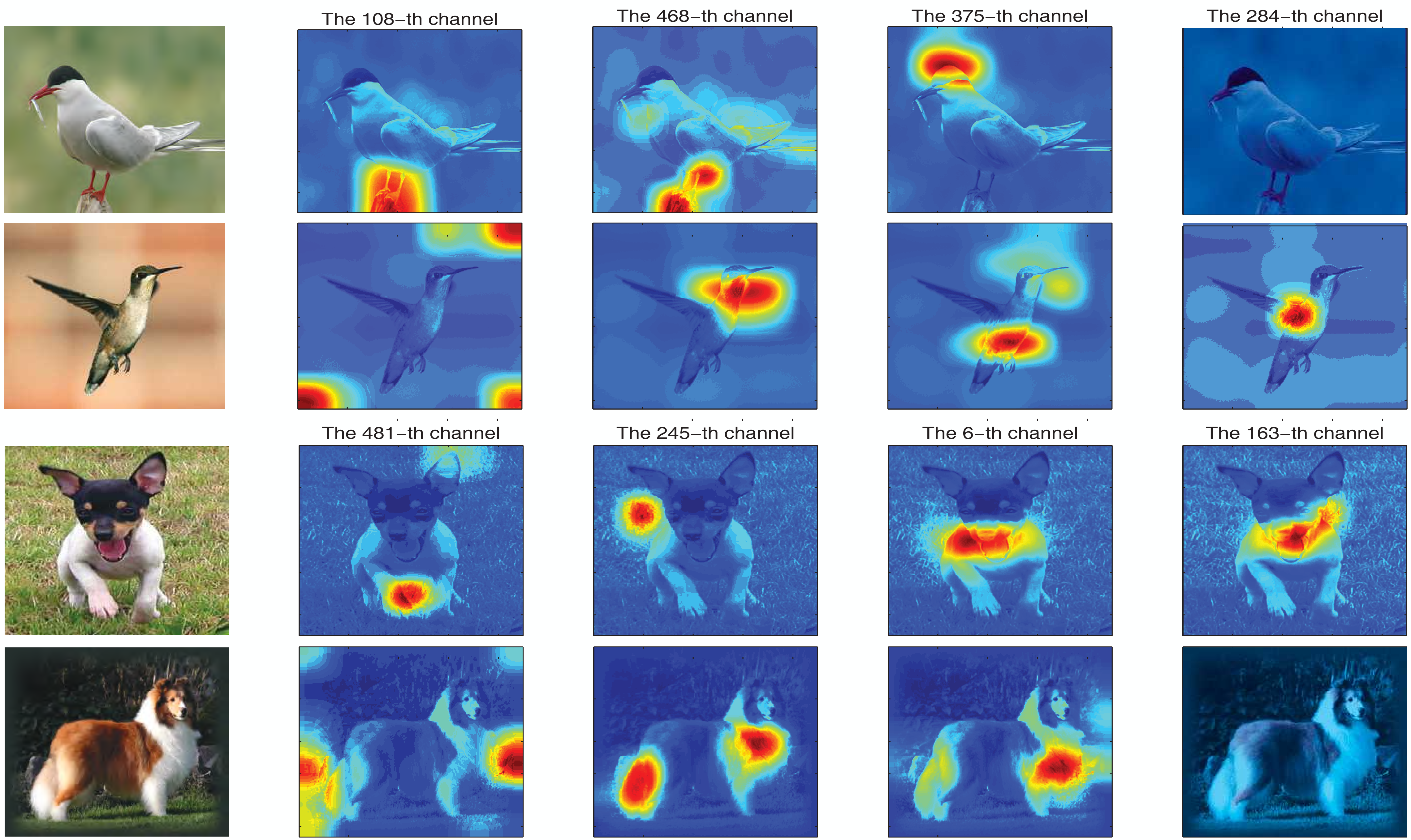
Feature maps:

2-D feature maps $S = \{S_n\}$ $(n = 1, \ldots, d)$

Descriptors:

$X = \{\boldsymbol{x}_{(i,j)}\}$

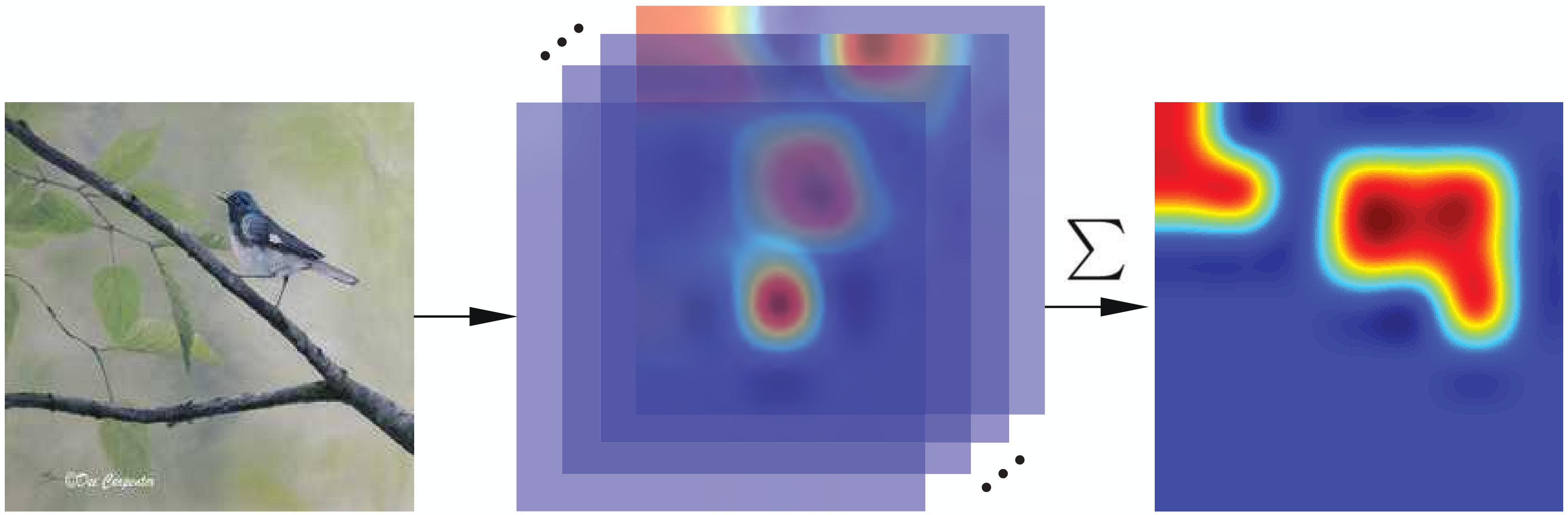
Fine-grained image retrieval (con't)



[Wei et al., IEEE TIP 2017]

Fine-grained image retrieval (con't)

Obtaining the activation map by summarizing feature maps



(a) Input image (b) Feature maps (c) Activation map

$$M_{i,j} = \begin{cases} 1 & \text{if } A_{i,j} > \bar{a} \\ 0 & \text{otherwise} \end{cases}$$

[Wei et al., IEEE TIP 2017]

Fine-grained image retrieval (con’t)

Visualization of the mask map M

(a) Visualization of the mask map M

(b) Visualization of the mask map $\widetilde{M}$

Fine-grained image retrieval (con't)

Selecting useful deep convolutional descriptors

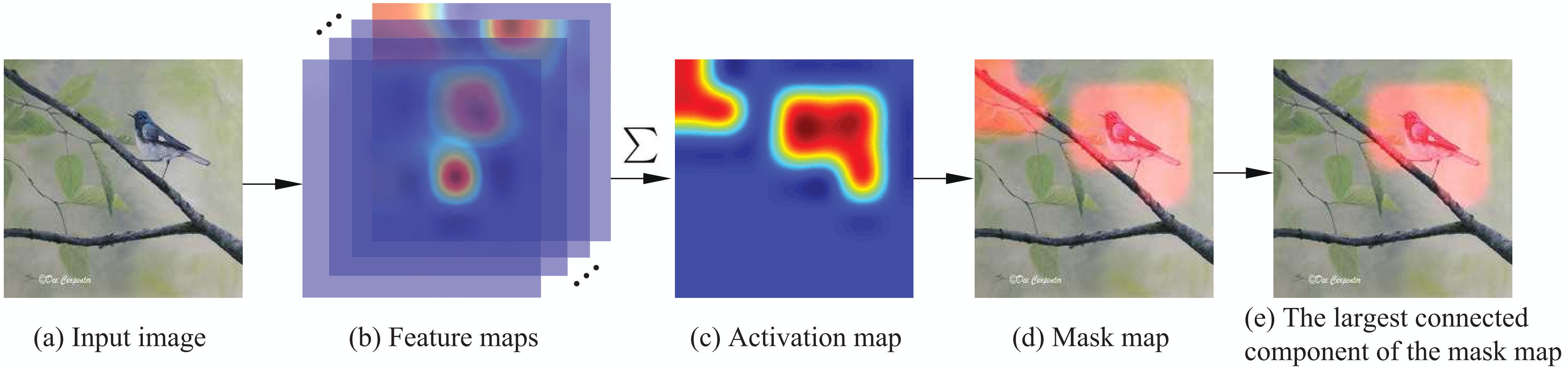
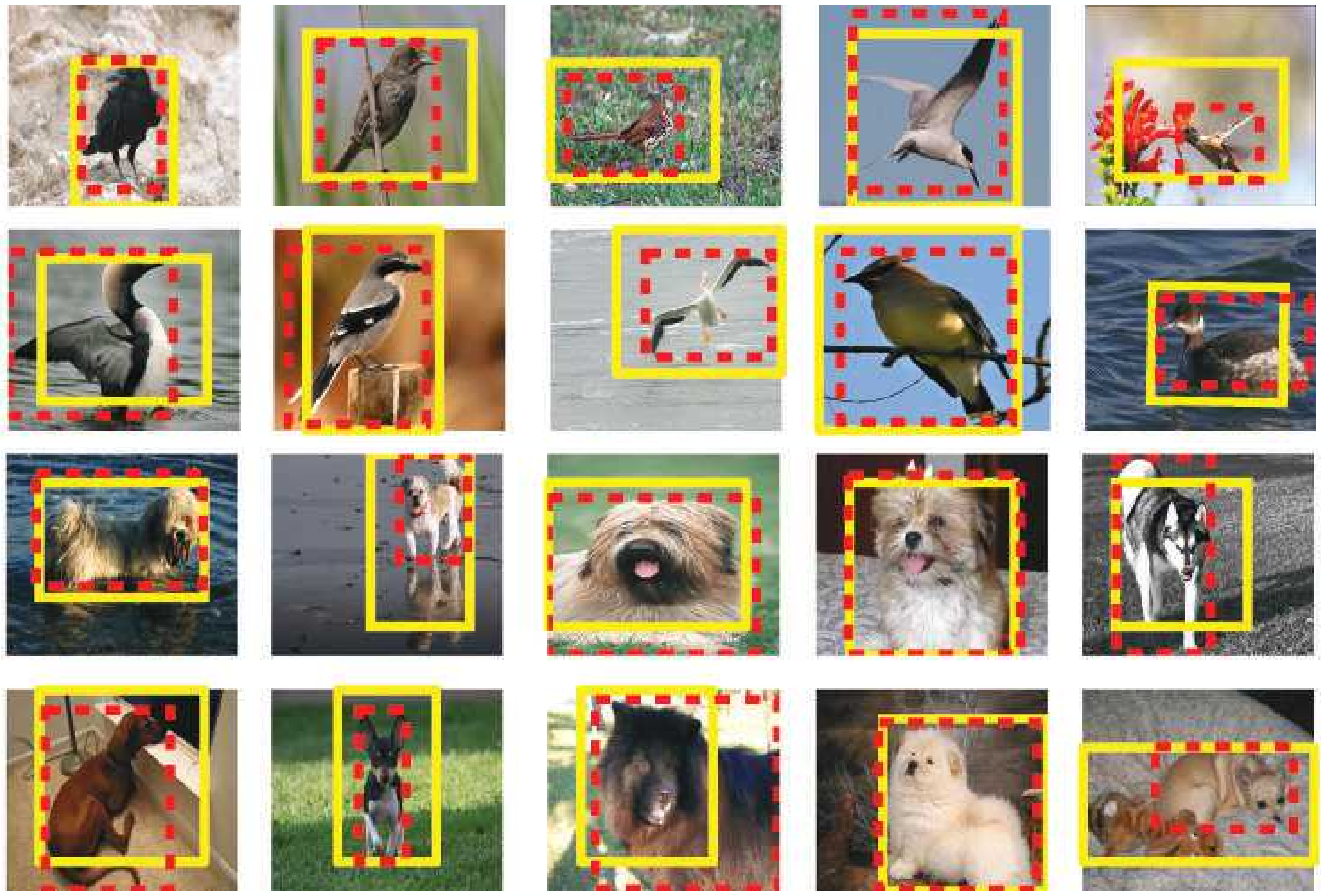


Figure 4. Selecting useful deep convolutional descriptors. (Best viewed in color.)

Fine-grained image retrieval (con't)

Qualitative evaluation



Fine-grained image retrieval (con't)

Aggregating convolutional descriptors

– **VLAD** [14] uses k-means to find a codebook of K centroids $\{\boldsymbol{c}_1, \dots, \boldsymbol{c}_K\}$ and maps $\boldsymbol{x}_{(i,j)}$ into a single vector $\boldsymbol{v}_{(i,j)} = \begin{bmatrix}\mathbf{0} & \dots & \mathbf{0} & \boldsymbol{x}_{(i,j)} - \boldsymbol{c}_k & \dots & \mathbf{0}\end{bmatrix} \in \mathcal{R}^{K \times d}$, where $\boldsymbol{c}_k$ is the closest centroid to $\boldsymbol{x}_{(i,j)}$. The final representation is $\sum_{i,j} \boldsymbol{v}_{(i,j)}$.

– **Fisher Vector** [15]: FV is similar to VLAD, but uses a soft assignment (i.e., Gaussian Mixture Model) instead of using k-means. Moreover, FV also includes second-order statistics.[2]

– **Pooling approaches**. We also try two traditional pooling approaches, i.e., max-pooling and average-pooling, to aggregate the deep descriptors.

Fine-grained image retrieval (con't)

Comparing difference encoding or pooling methods

Approach	Dimension	*CUB200-2011*		*Stanford Dogs*	
		top1	top5	top1	top5
VLAD	1,024	55.92%	62.51%	69.28%	74.43%
Fisher Vector	2,048	52.04%	59.19%	68.37%	73.74%
avgPool	512	56.42%	63.14%	73.76%	78.47%
maxPool	512	58.35%	64.18%	70.37%	75.59%
avg&maxPool	1,024	**59.72%**	**65.79%**	**74.86%**	**79.24%**

SCDA

Fine-grained image retrieval (con't)

Multiple layer ensemble

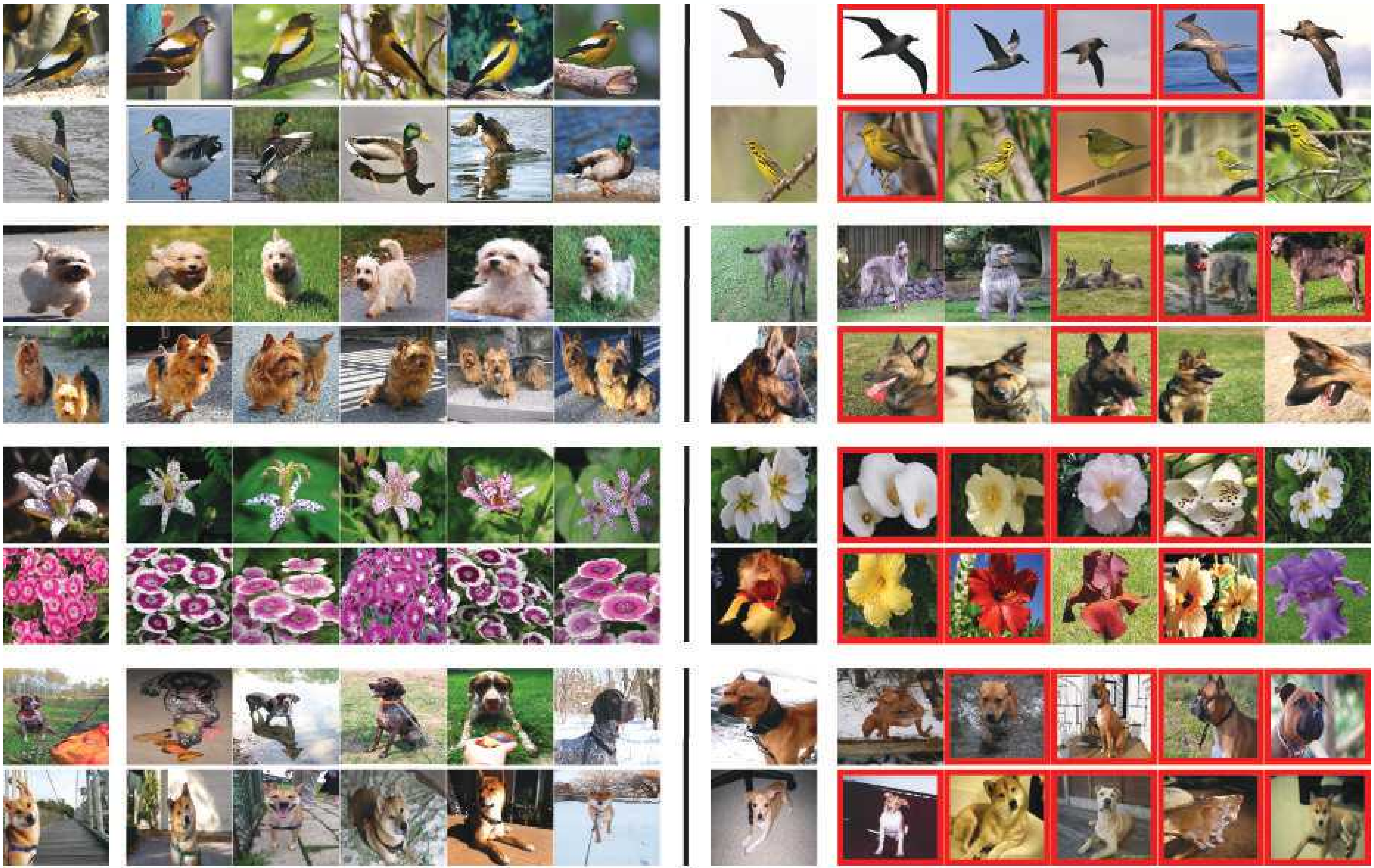
(a) M of Pool5 (b) $\widetilde{M}$ of Pool5 (c) M of Relu5_2 (d) $\widetilde{M}$ of Relu5_2

Figure 6. The mask map and its corresponding largest connected component of different CNN layers. (The figure is best viewed in color.)

$$\mathrm{SCDA}^{+} \leftarrow \left[\mathrm{SCDA}_{\mathrm{pool}_5},\ \alpha \times \mathrm{SCDA}_{\mathrm{relu}_{5\_2}}\right]$$

$$\mathrm{SCDA\_flip}^{+}$$

Fine-grained image retrieval (con't)



[Wei et al., IEEE TIP 2017]

http://www.weixiushen.com/

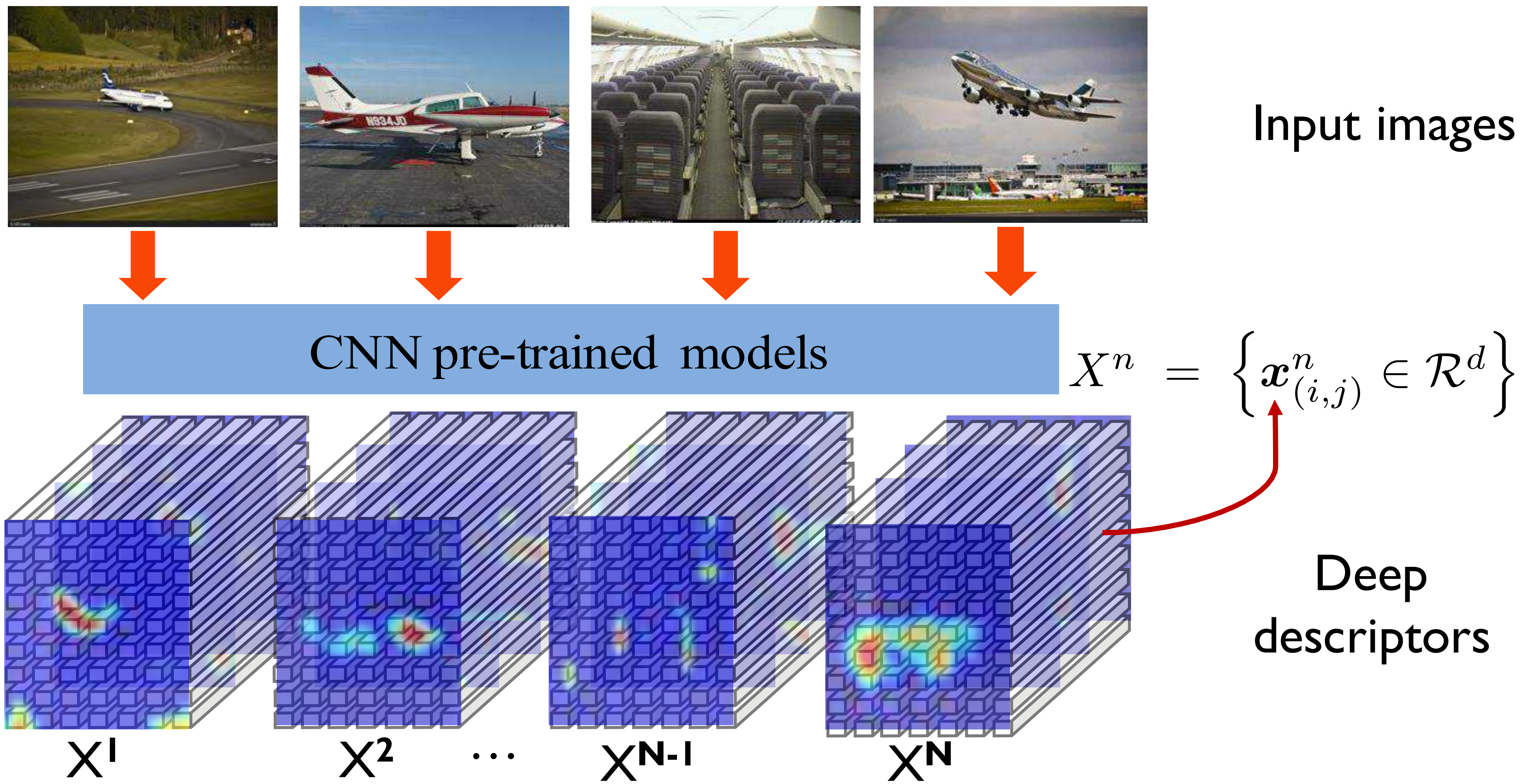
Fine-grained image retrieval (con’t)

Quality demonstration of the SCDA feature

[Wei et al., IEEE TIP 2017]

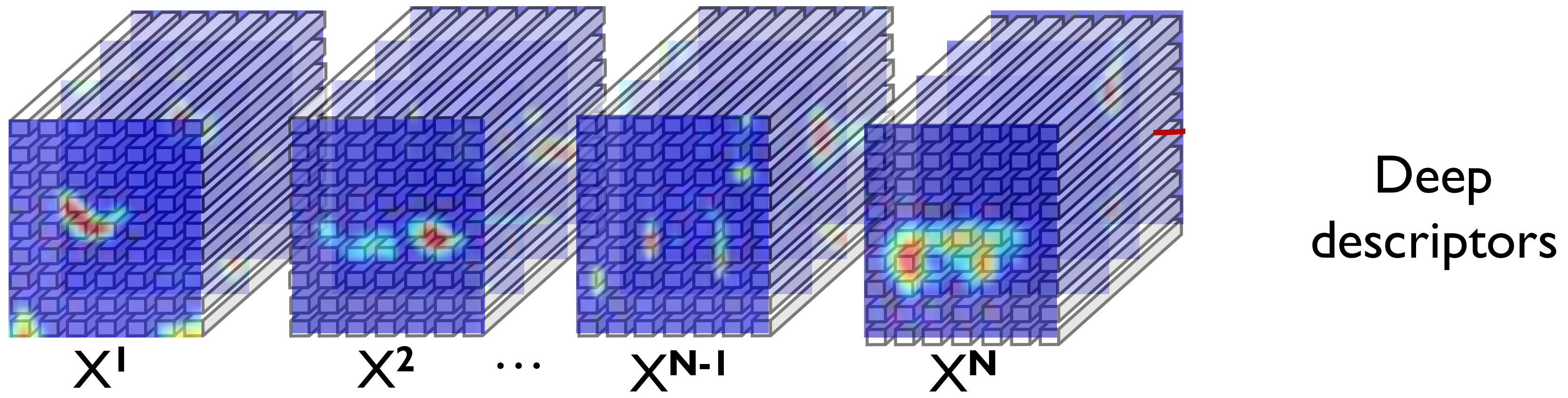
Fine-grained image retrieval (con't)

Deep Descriptor Transforming



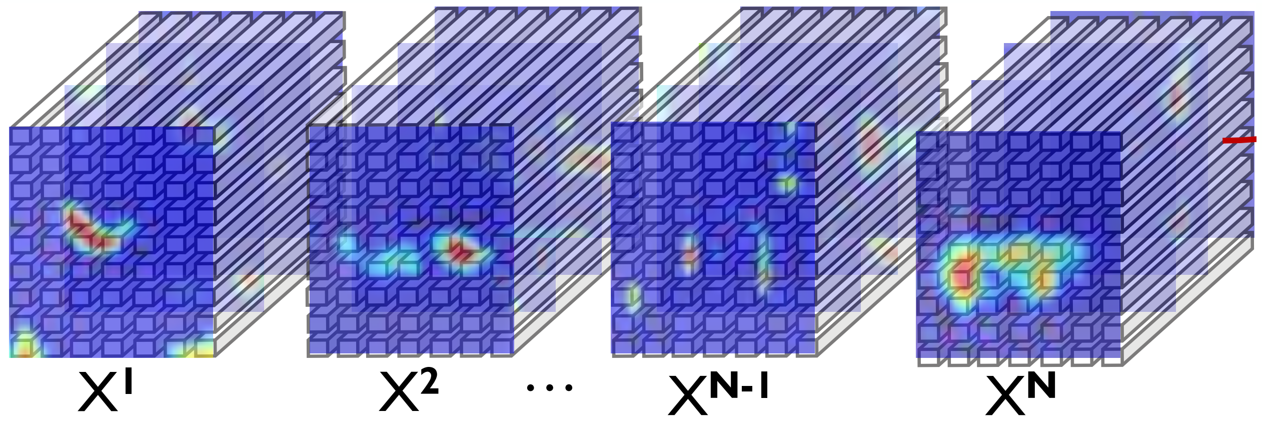
Fine-grained image retrieval (con't)

What we need is a mapping function …



Fine-grained image retrieval (con't)

What we need is a mapping function …



Deep descriptors

$$\bar{\boldsymbol{x}} = \frac{1}{K}\sum_{n}\sum_{i,j}\boldsymbol{x}^{n}_{(i,j)}$$

Calculating descriptors' mean vector

$$\mathrm{Cov}(\boldsymbol{x}) = \frac{1}{K}\sum_{n}\sum_{i,j}(\boldsymbol{x}^{n}_{(i,j)} - \bar{\boldsymbol{x}})(\boldsymbol{x}^{n}_{(i,j)} - \bar{\boldsymbol{x}})^{\top}$$

$$\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_d \text{ of } \mathrm{Cov}(\boldsymbol{x})$$

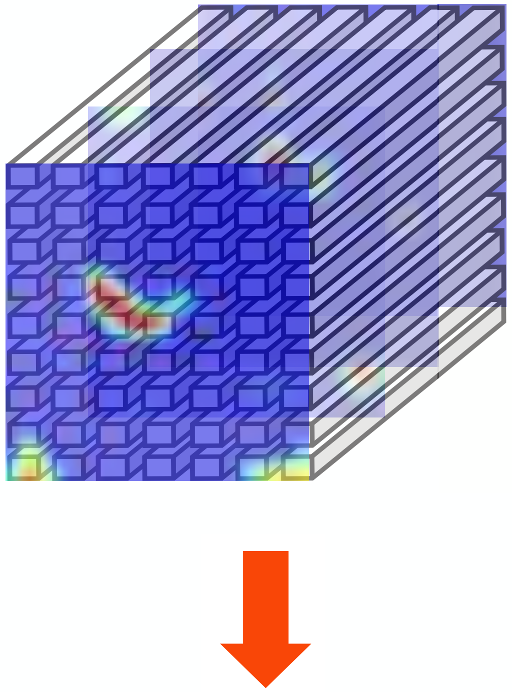
Obtaining the covariance matrix and then getting the eigenvectors

$$p^{1}_{(i,j)} = \boldsymbol{\xi}_1^{\top}\left(\boldsymbol{x}_{(i,j)} - \bar{\boldsymbol{x}}\right)$$

Transforming

Fine-grained image retrieval (con't)

Use the first eigenvector by PCA as the projection direction

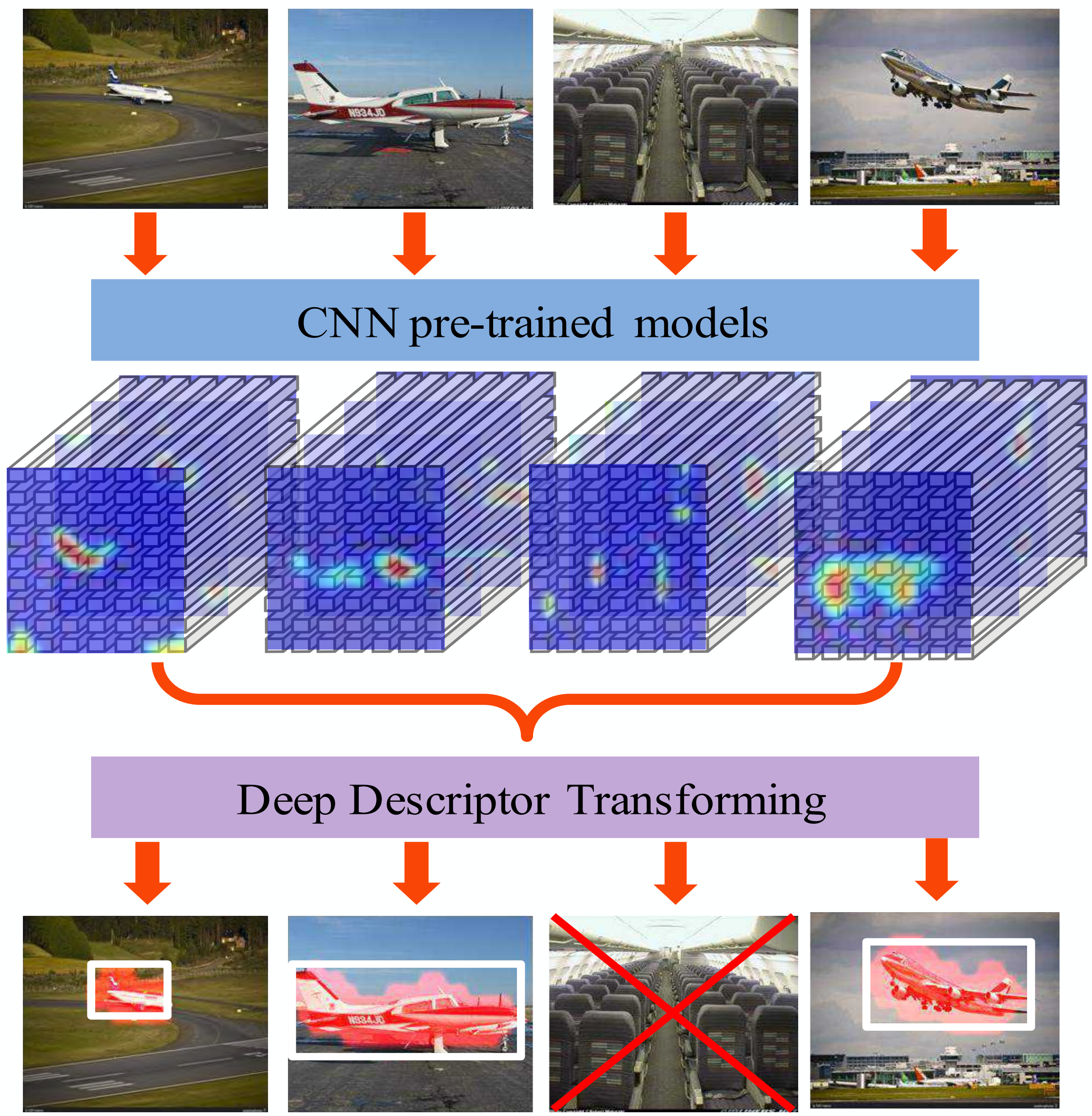


$$P^1 = \begin{bmatrix} p^1_{(1,1)} & p^1_{(1,2)} & \cdots & p^1_{(1,w)} \\ p^1_{(2,1)} & p^1_{(2,2)} & \cdots & p^1_{(2,w)} \\ \vdots & \vdots & \ddots & \vdots \\ p^1_{(h,1)} & p^1_{(h,2)} & \cdots & p^1_{(h,w)} \end{bmatrix}$$

Indicator matrix for co-localization

Fine-grained image retrieval (con't)

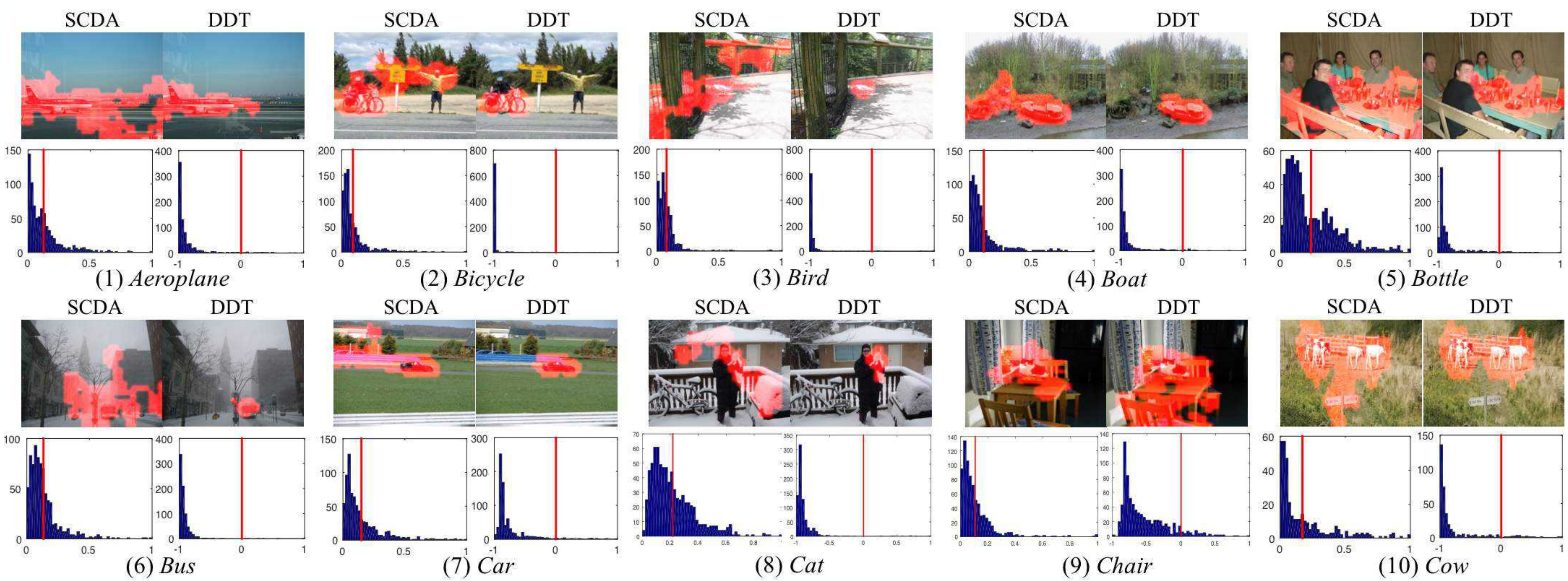
The whole pipeline of DDT



[Wei et al., IJCAI 2017]

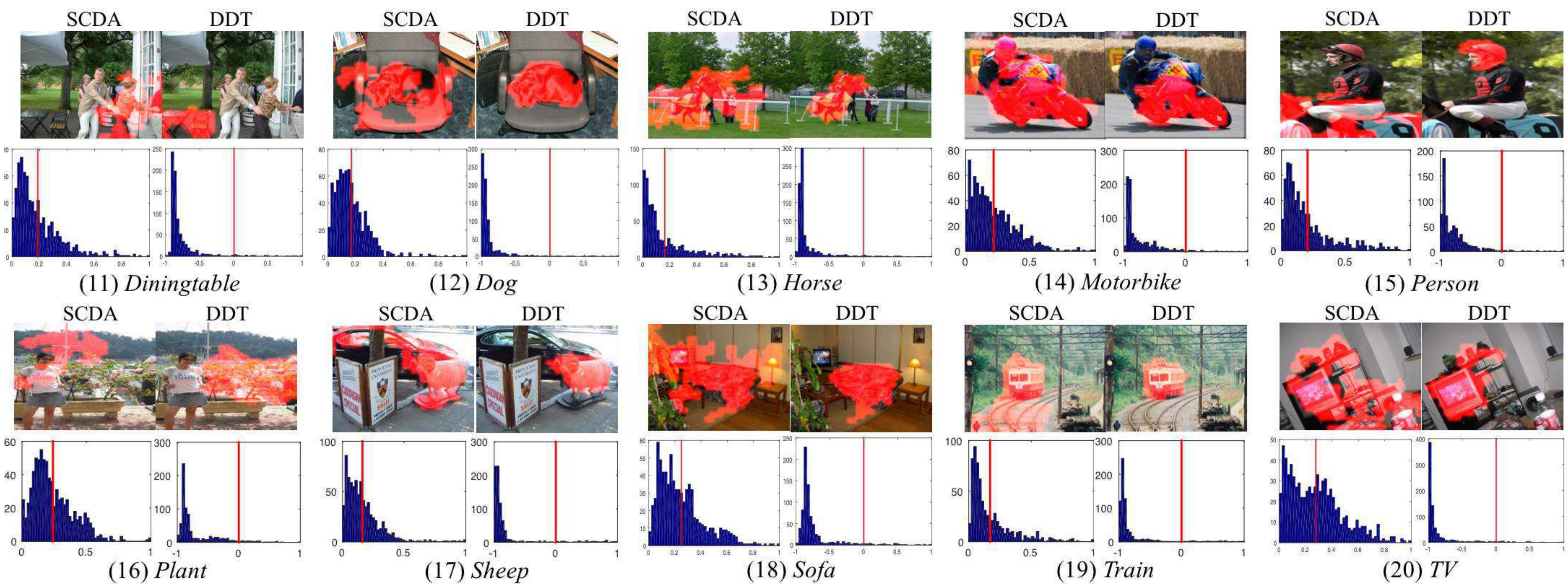
Fine-grained image retrieval (con't)

DDT *vs.* SCDA



Fine-grained image retrieval (con't)

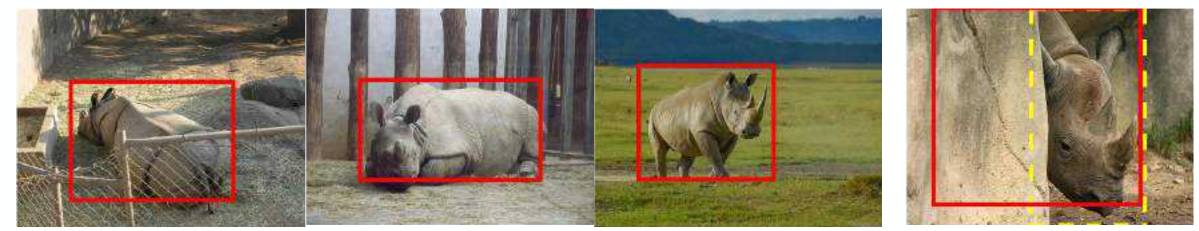
DDT *vs.* SCDA



Fine-grained image retrieval (con't)

Empirical results on *ImageNet-Subset (disjoint with ImageNet)*

(a) *Chipmunk*



(b) *Rhino*

(c) *Stoat*

(d) *Racoon*

(e) *Rake*

(f) *Wheelchair*

[Wei et al., IJCAI 2017]

http://www.weixiushen.com/

Fine-grained image retrieval (con't)

Extension to video co-localization

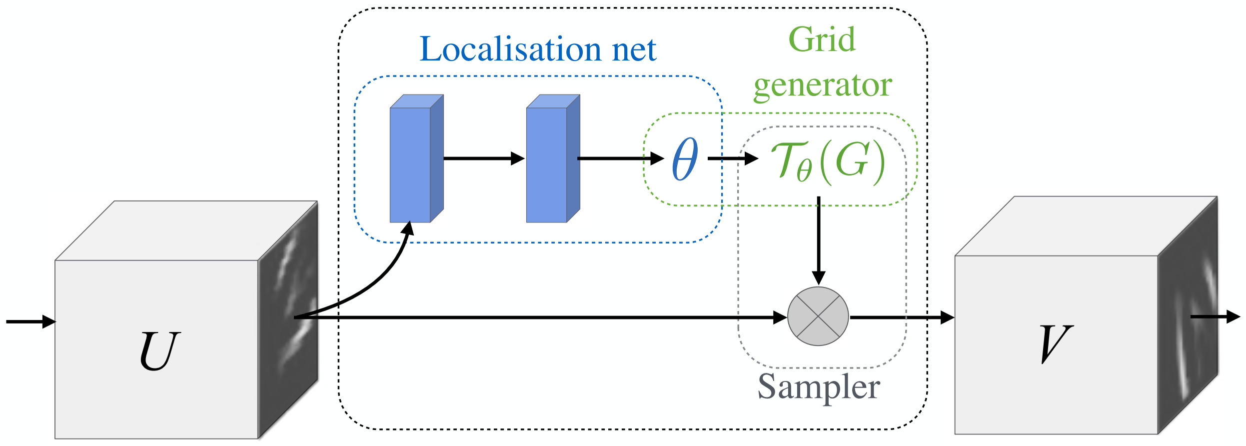
Part 2

- Fine-grained image recognition
  - Fine-grained image recognition with powerful representation learning
  - Fine-grained image recognition with part-based approaches
- Other computer vision tasks related to fine-grained image analysis
  - Person / Vehicle re-identification
  - Clothes retrieval
  - Product recognition
- New developments of fine-grained image analysis
  - Fine-grained images with languages
  - Few-shot fine-grained image recognition

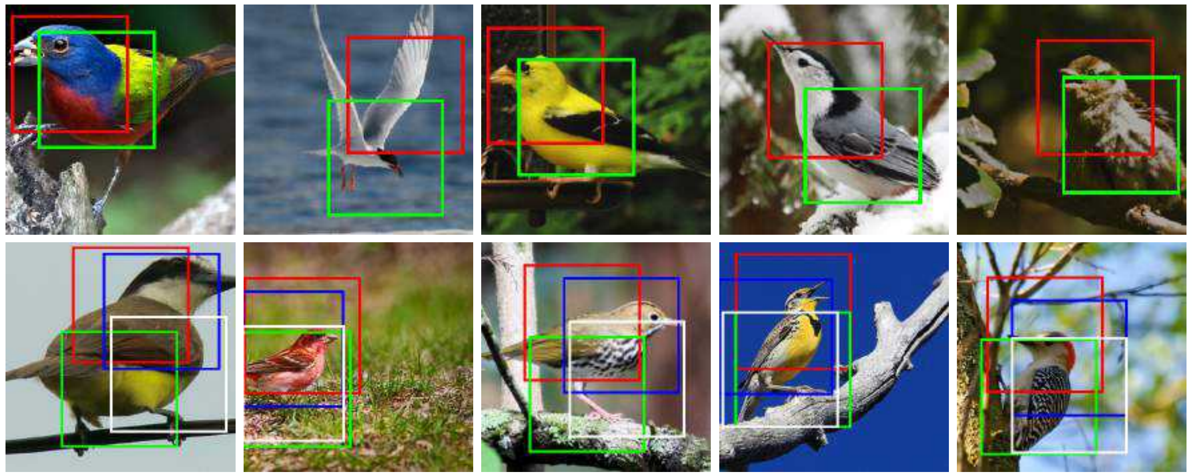
Fine-grained image recognition

Fine-grained image recognition with end-to-end feature encoding

Spatial Transformer Networks

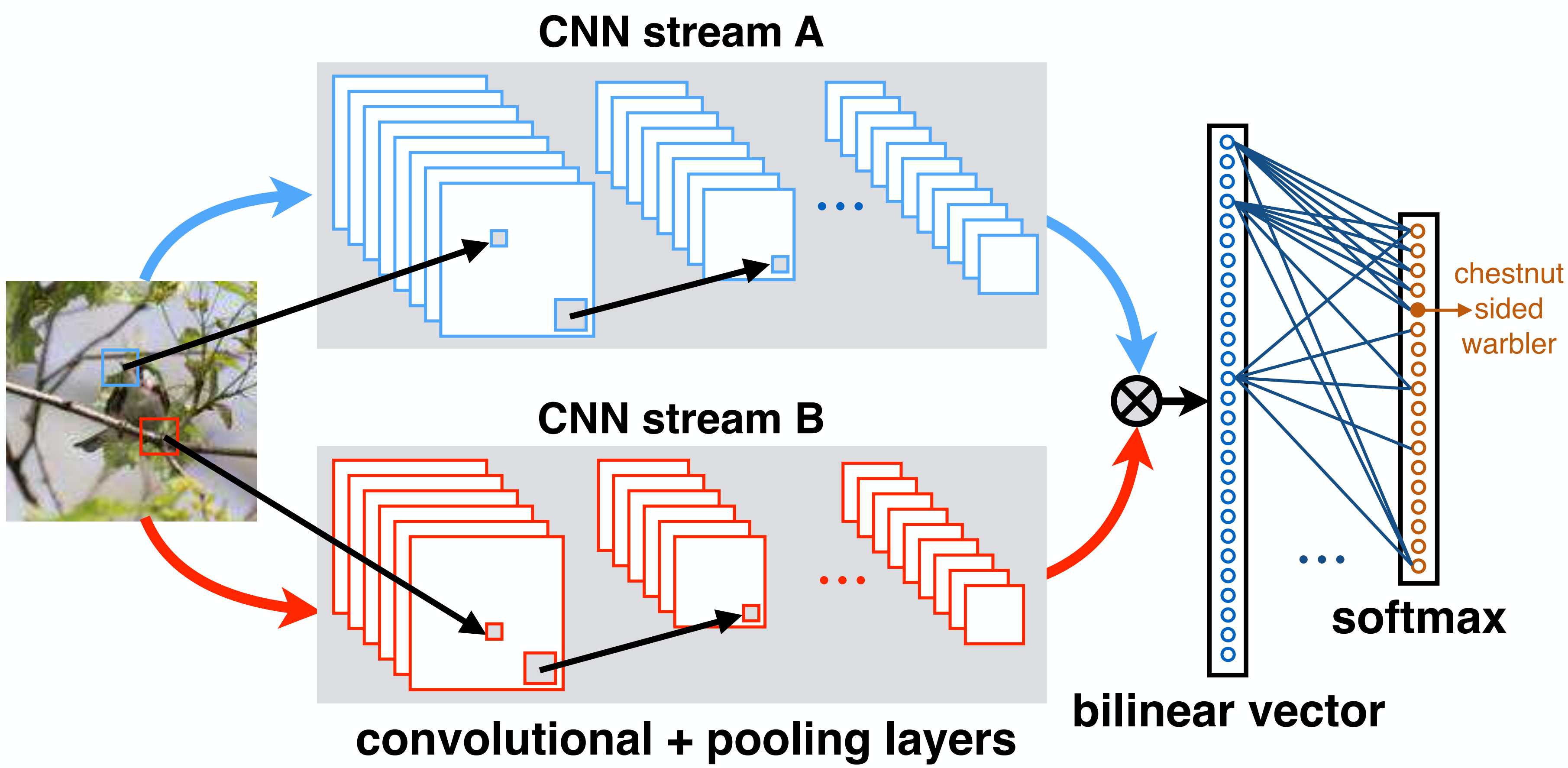


Model		
Cimpoi '15 [4]		66.7
Zhang '14 [30]		74.9
Branson '14 [2]		75.7
Lin '15 [20]		80.9
Simon '15 [24]		81.0
CNN (ours)	224px	82.3
2×ST-CNN	224px	83.1
2×ST-CNN	448px	83.9
4×ST-CNN	448px	**84.1**



Fine-grained image recognition (con't)

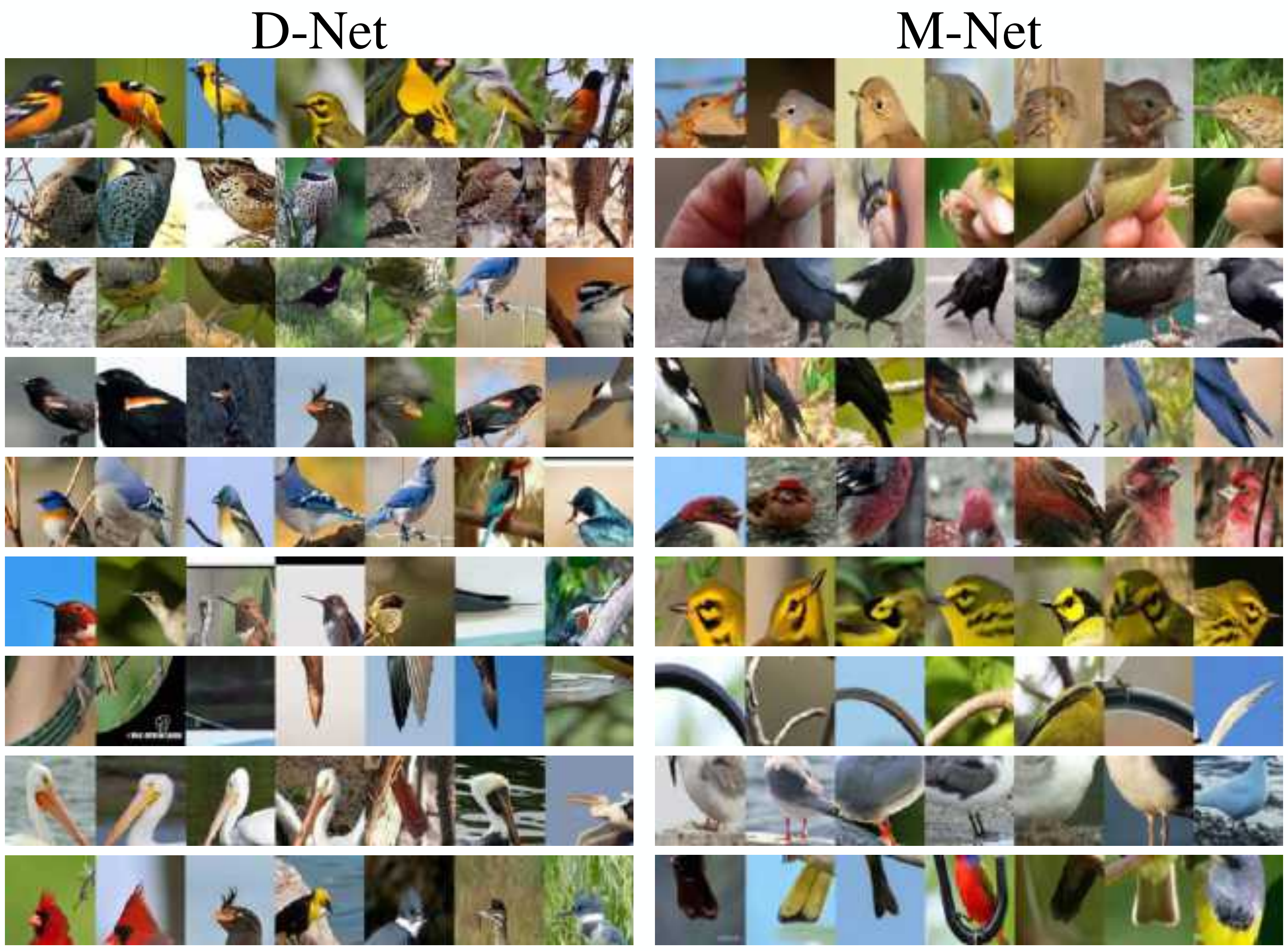
Fine-grained image recognition with end-to-end feature encoding



Bilinear Convolutional Neural Networks

Fine-grained image recognition (con't)

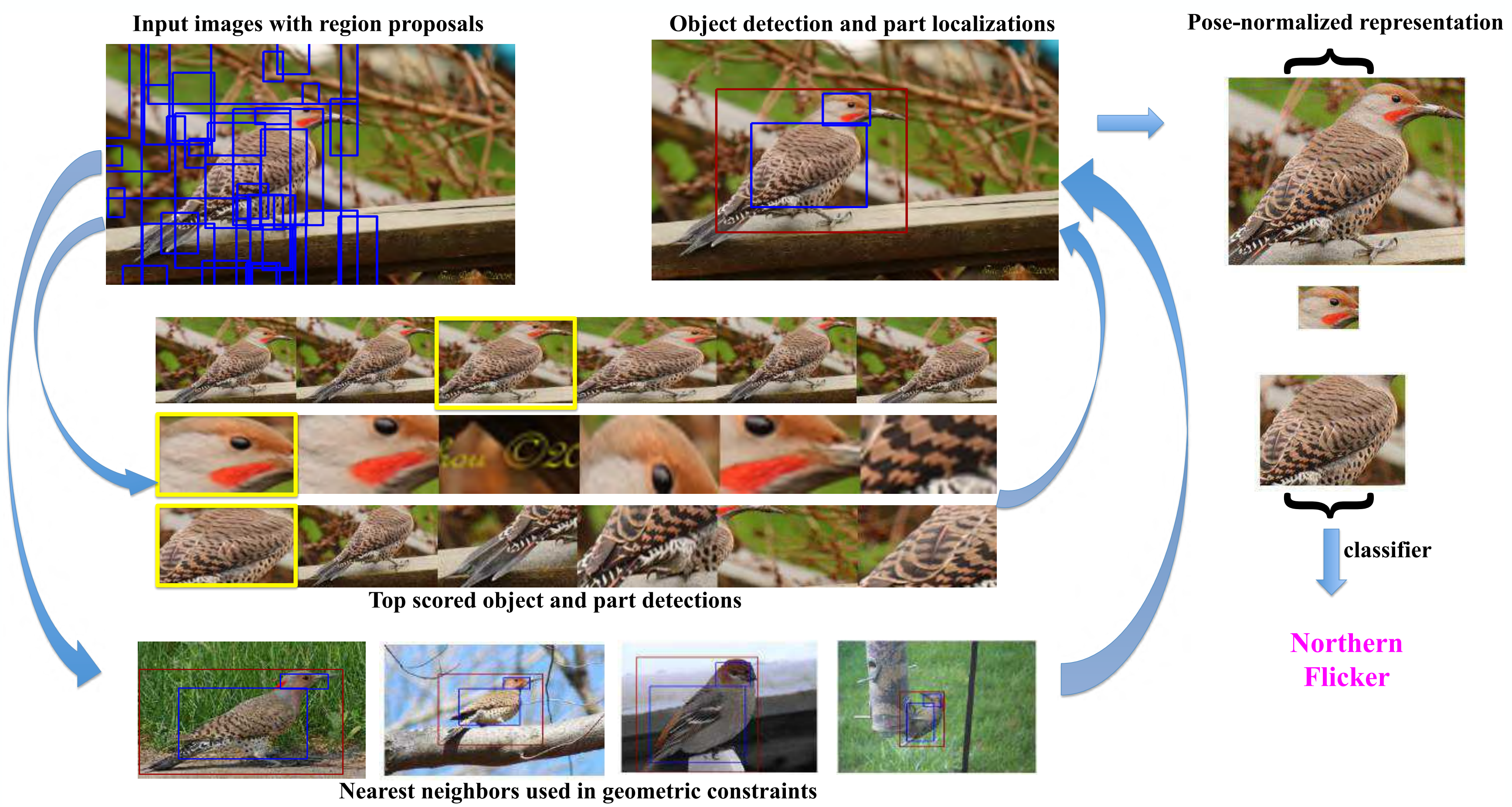
Fine-grained image recognition with end-to-end feature encoding



Qualitative results of Bilinear CNNs

Fine-grained image recognition (con't)

Fine-grained image recognition by localization-classification subnetworks

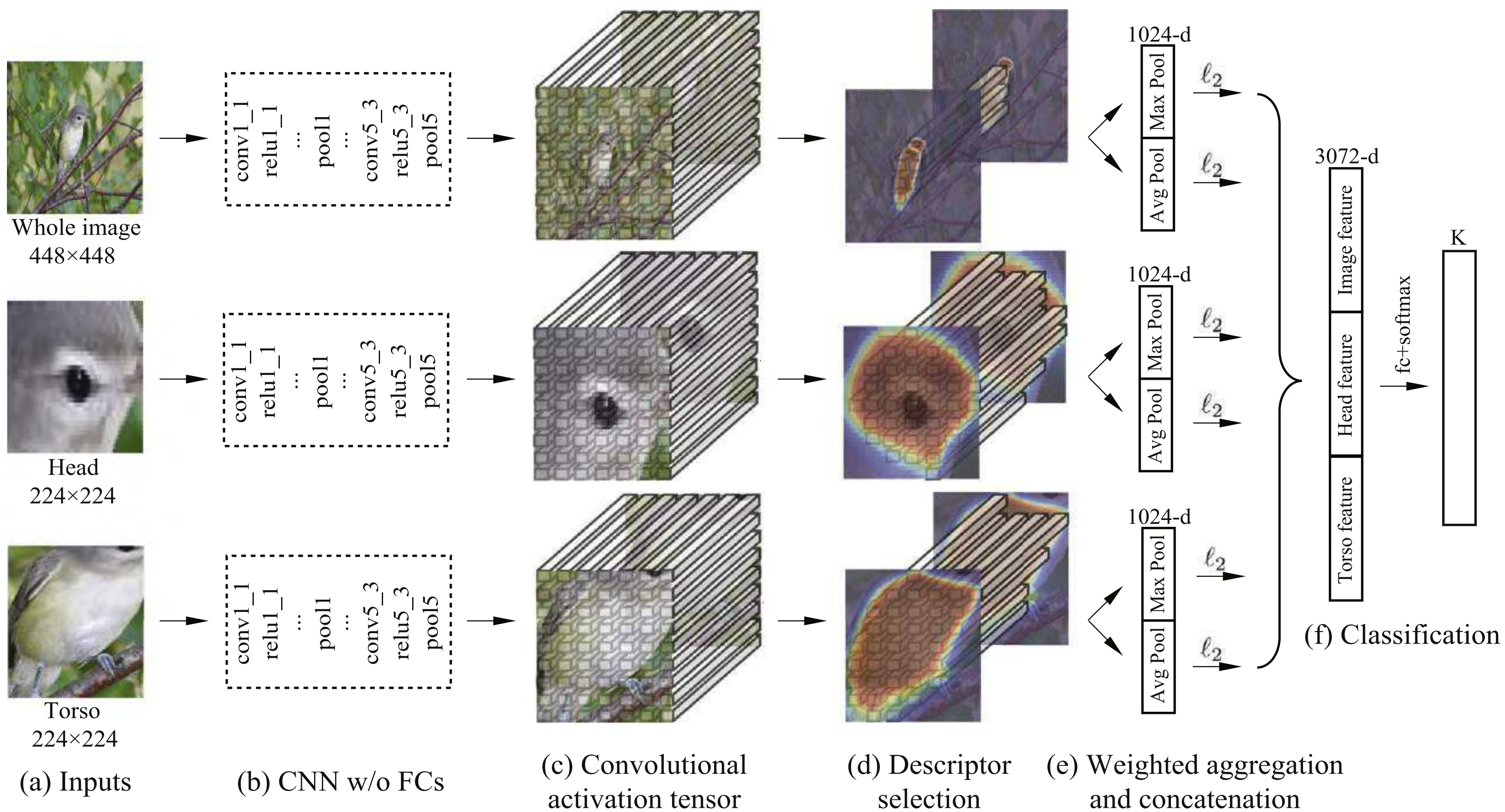


Part-based R-CNNs

[Zhang et al., ECCV 2014]

Fine-grained image recognition (con't)

Fine-grained image recognition by localization-classification subnetworks



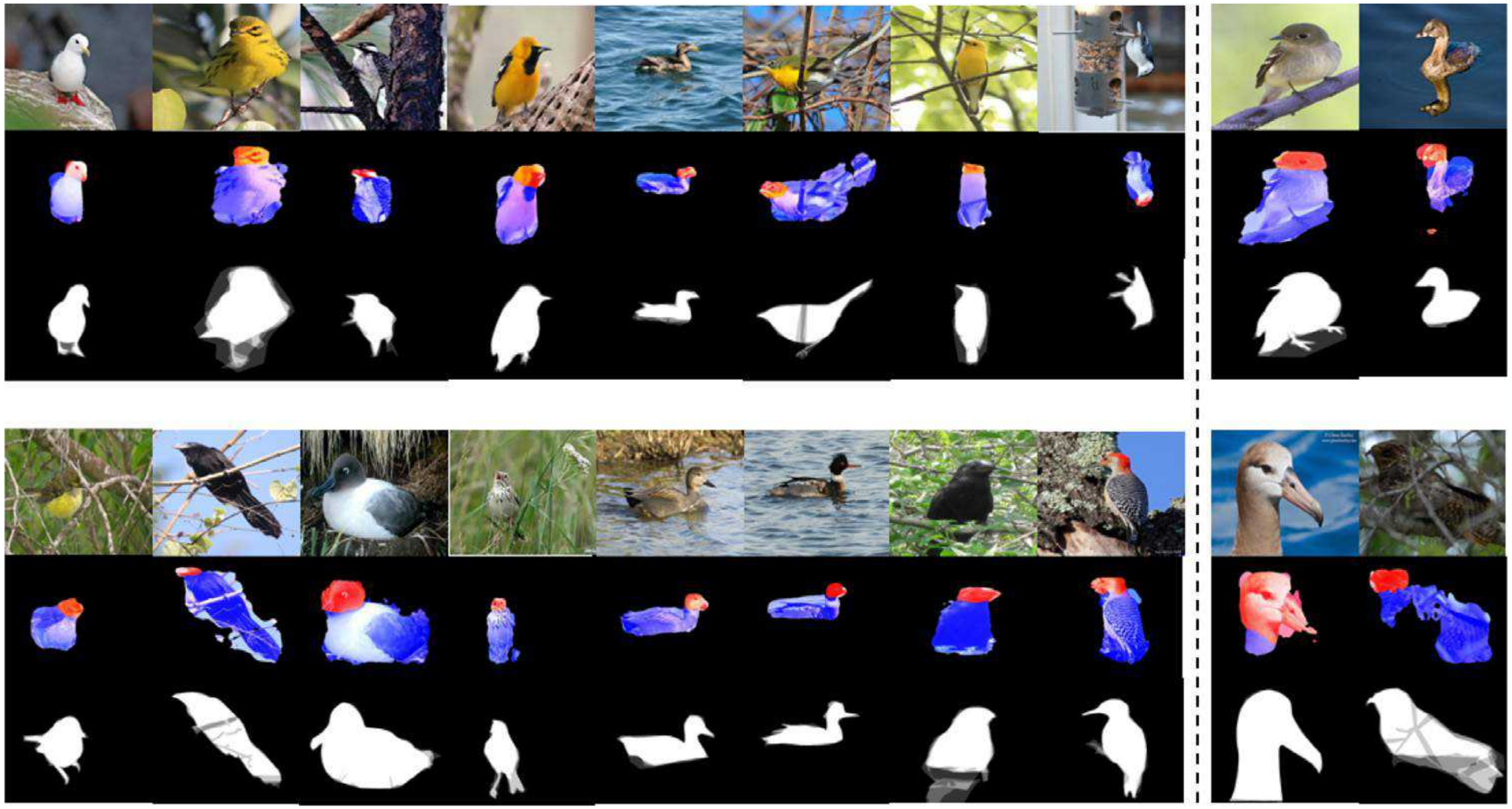
Mask-CNN

[Wei et al., Pattern Recognition 2018]

http://www.weixiushen.com/

Fine-grained image recognition (con’t)

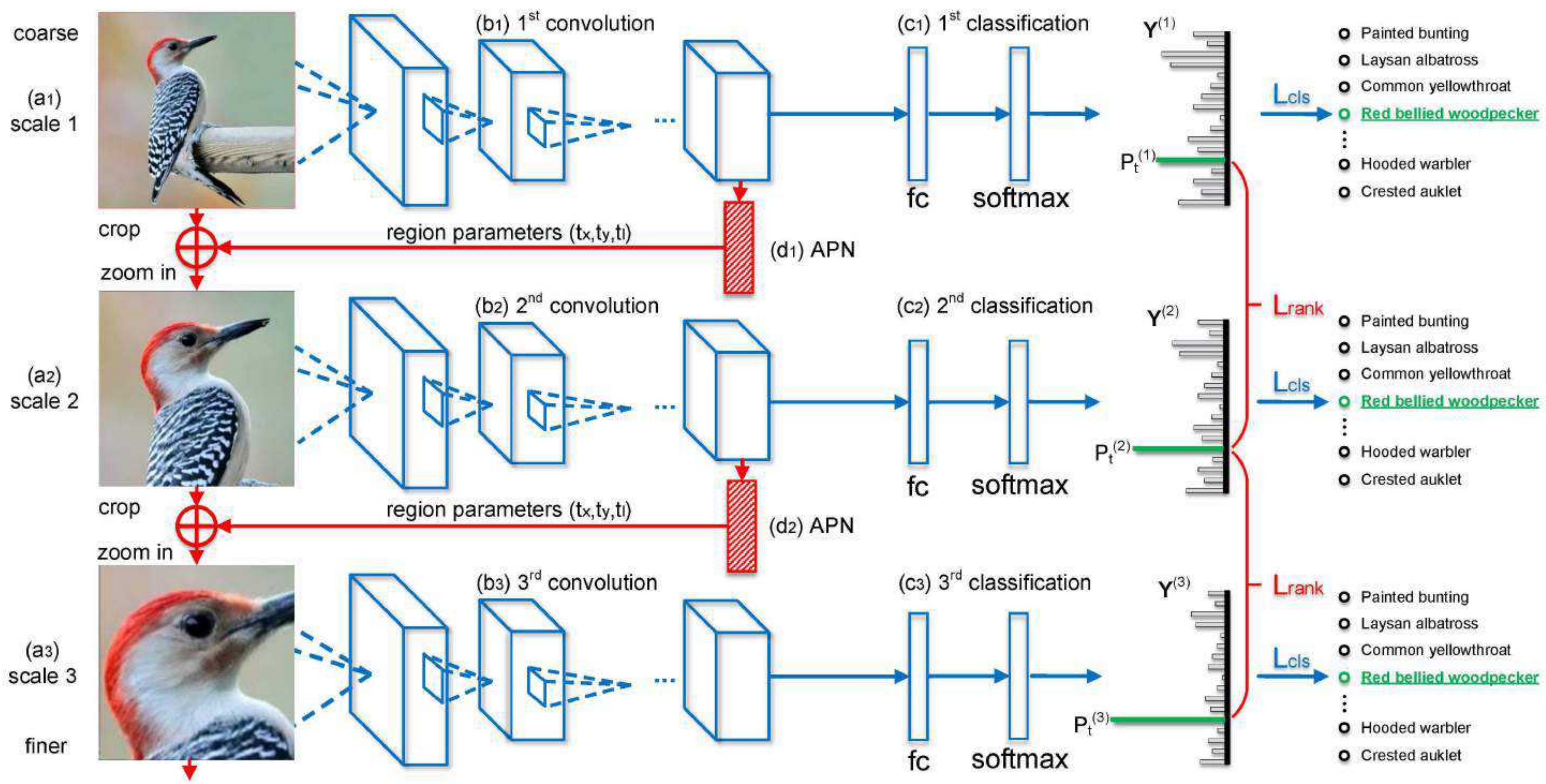
Fine-grained image recognition by localization-classification subnetworks



Qualitative results of Mask-CNN

Fine-grained image recognition (con't)

Fine-grained image recognition by localization-classification subnetworks



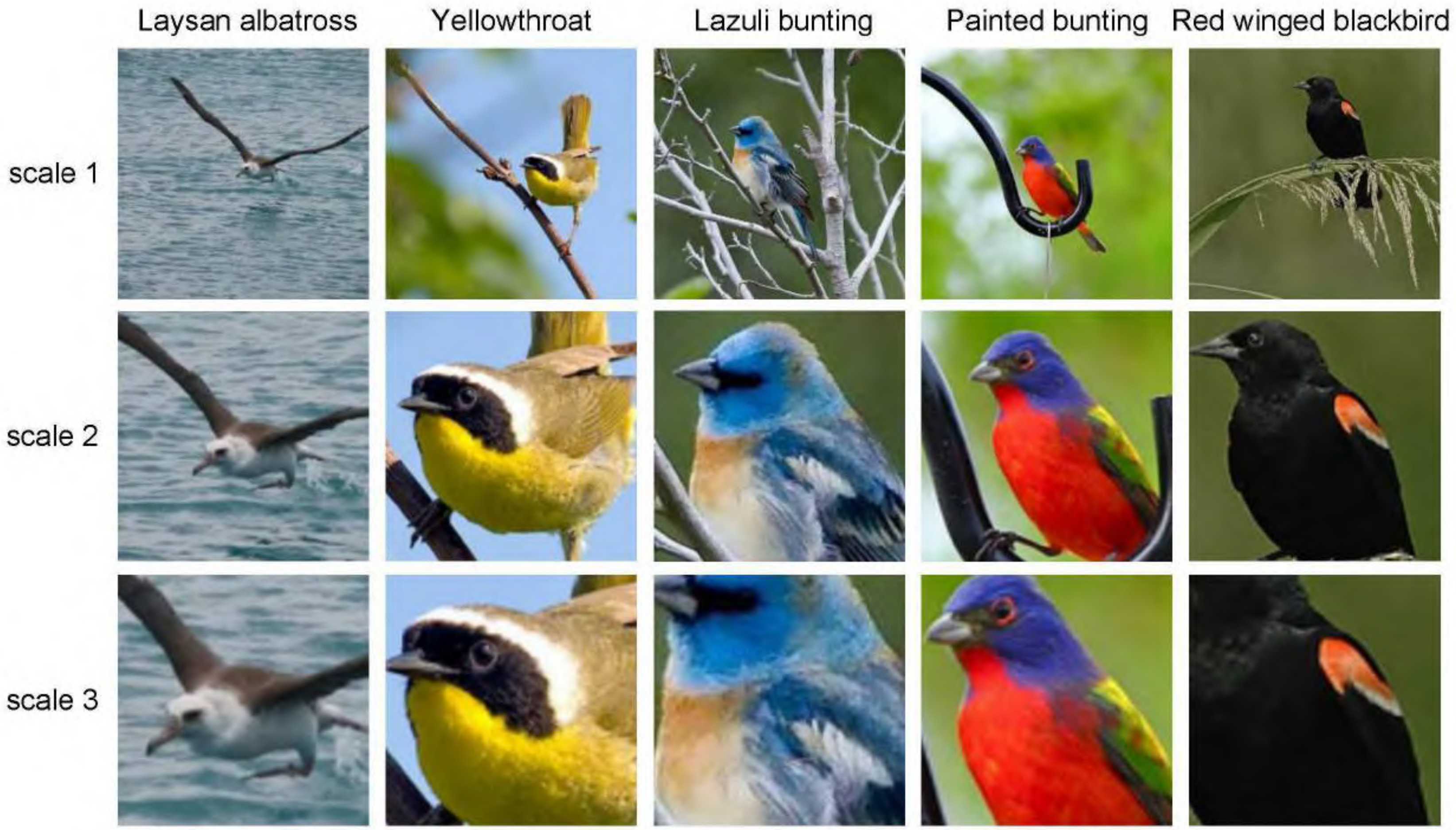
Recurrent CNN (RA-CNN)

[Fu et al., CVPR 2017]

http://www.weixiushen.com/

Fine-grained image recognition (con’t)

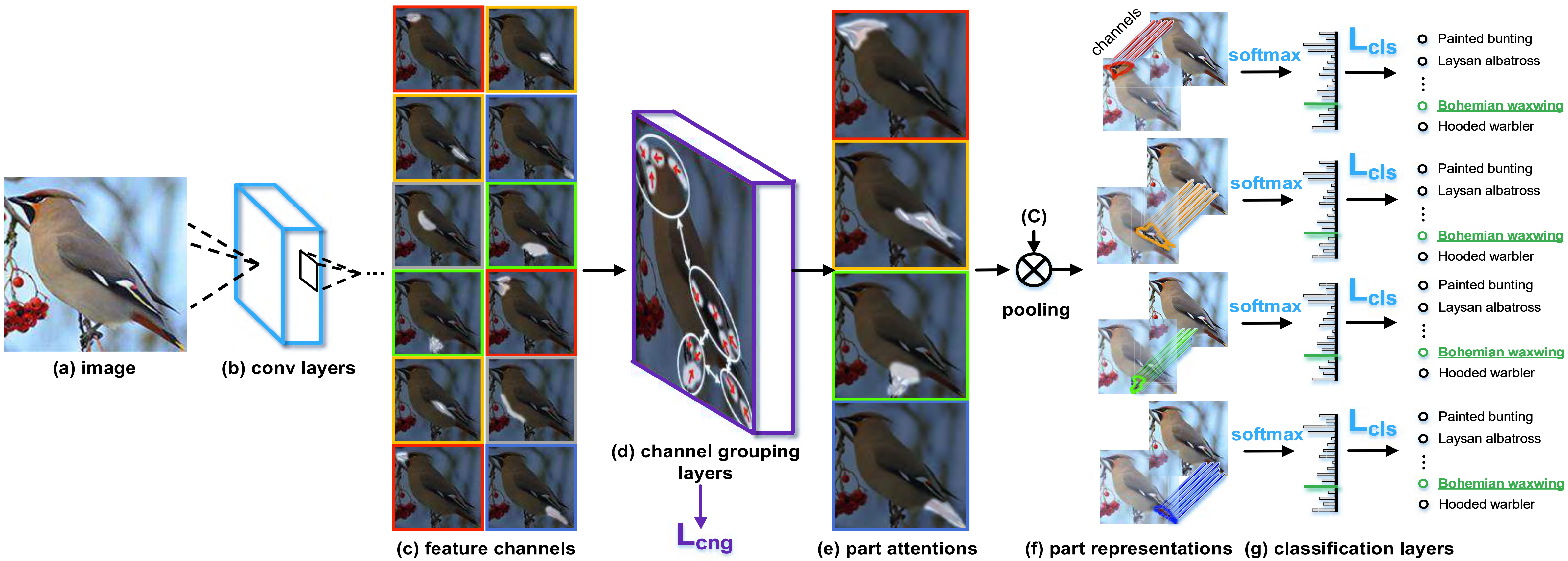
Fine-grained image recognition by localization-classification subnetworks



Qualitative results of RA-CNN

Fine-grained image recognition (con't)

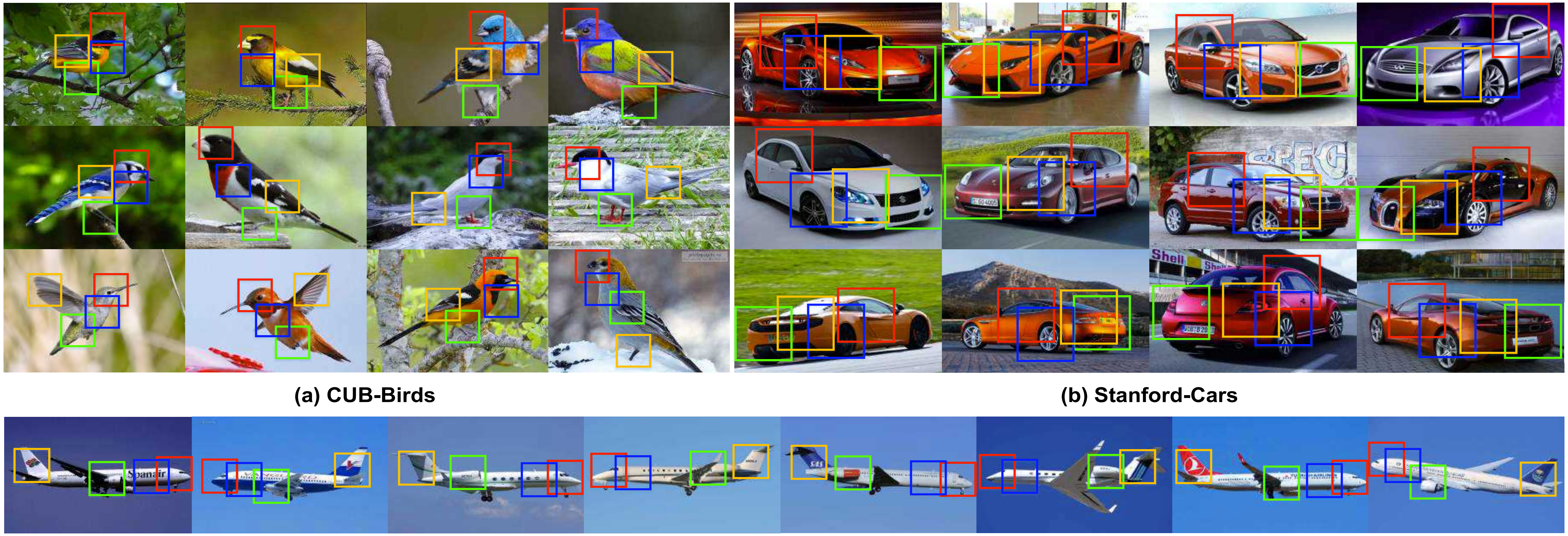
Fine-grained image recognition by localization-classification subnetworks



Multiple attention CNNs (MA-CNN)

Fine-grained image recognition (con't)

Fine-grained image recognition by localization-classification subnetworks



(a) CUB-Birds

(b) Stanford-Cars

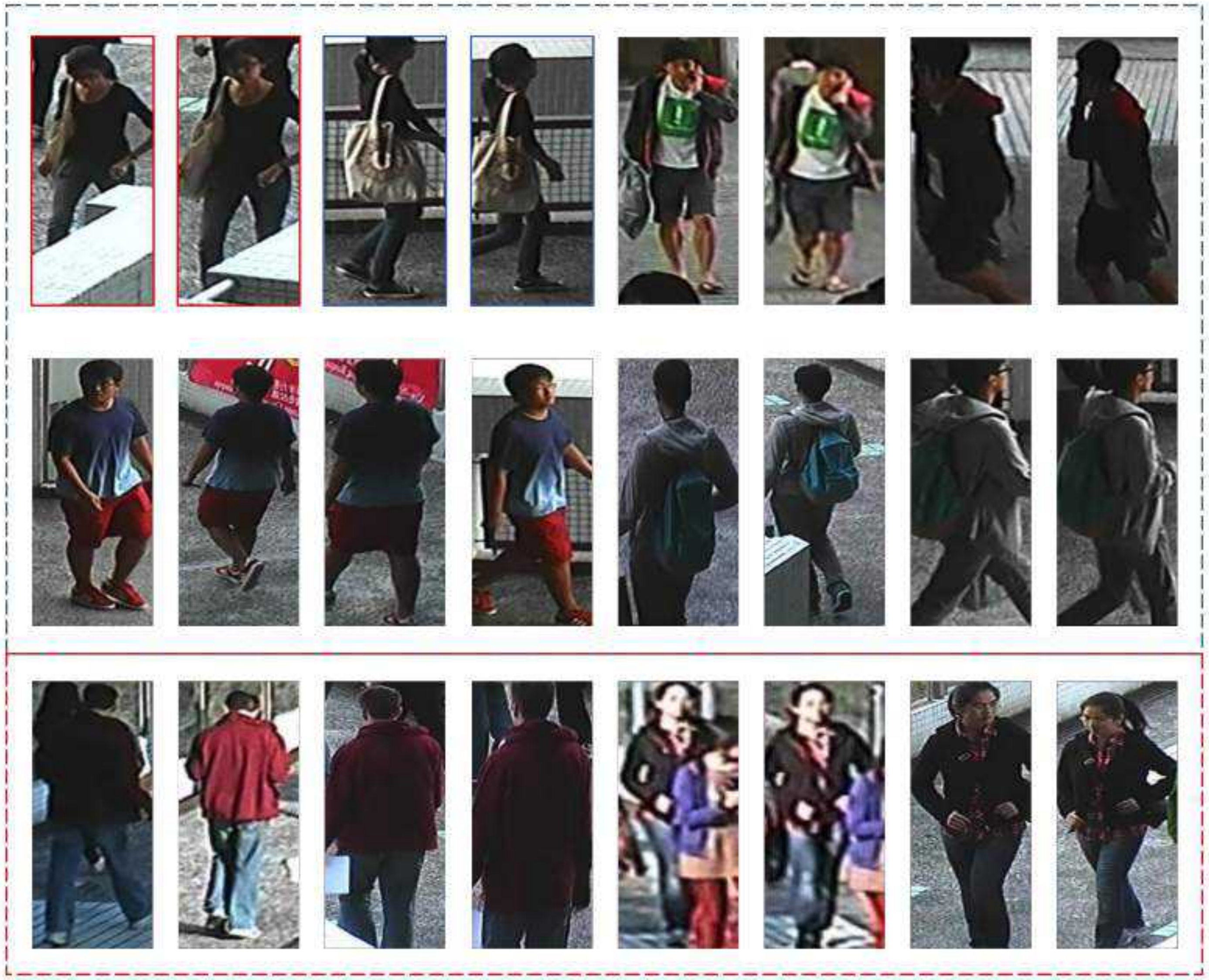
(c) FGVC-Aircraft

Qualitative results of MA-CNN

[Zheng et al., ICCV 2017]

Other CV tasks related to fine-grained

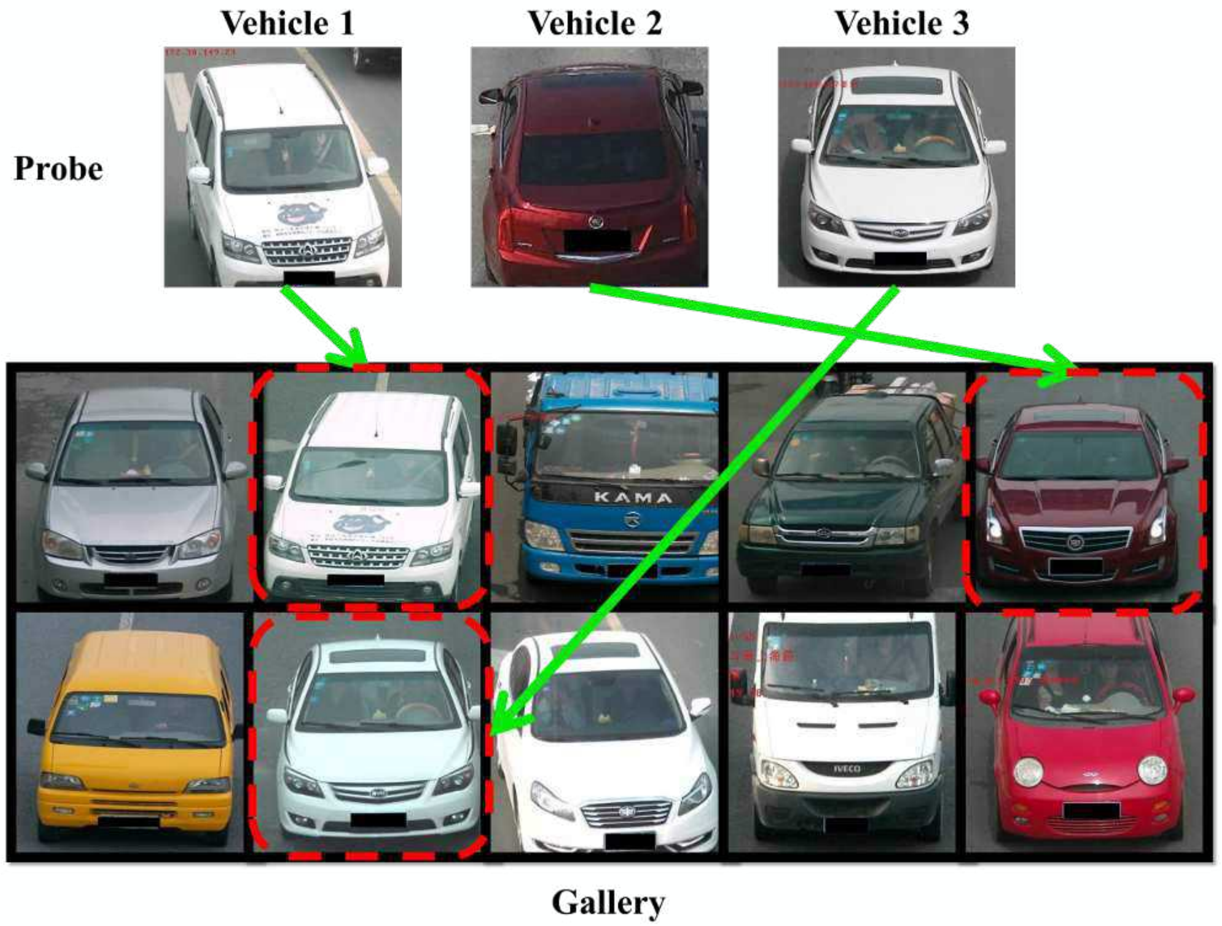
Person re-identification



Figures are courtesy of [Li et al., CVPR 2014].

Other CV tasks related to fine-grained (con't)

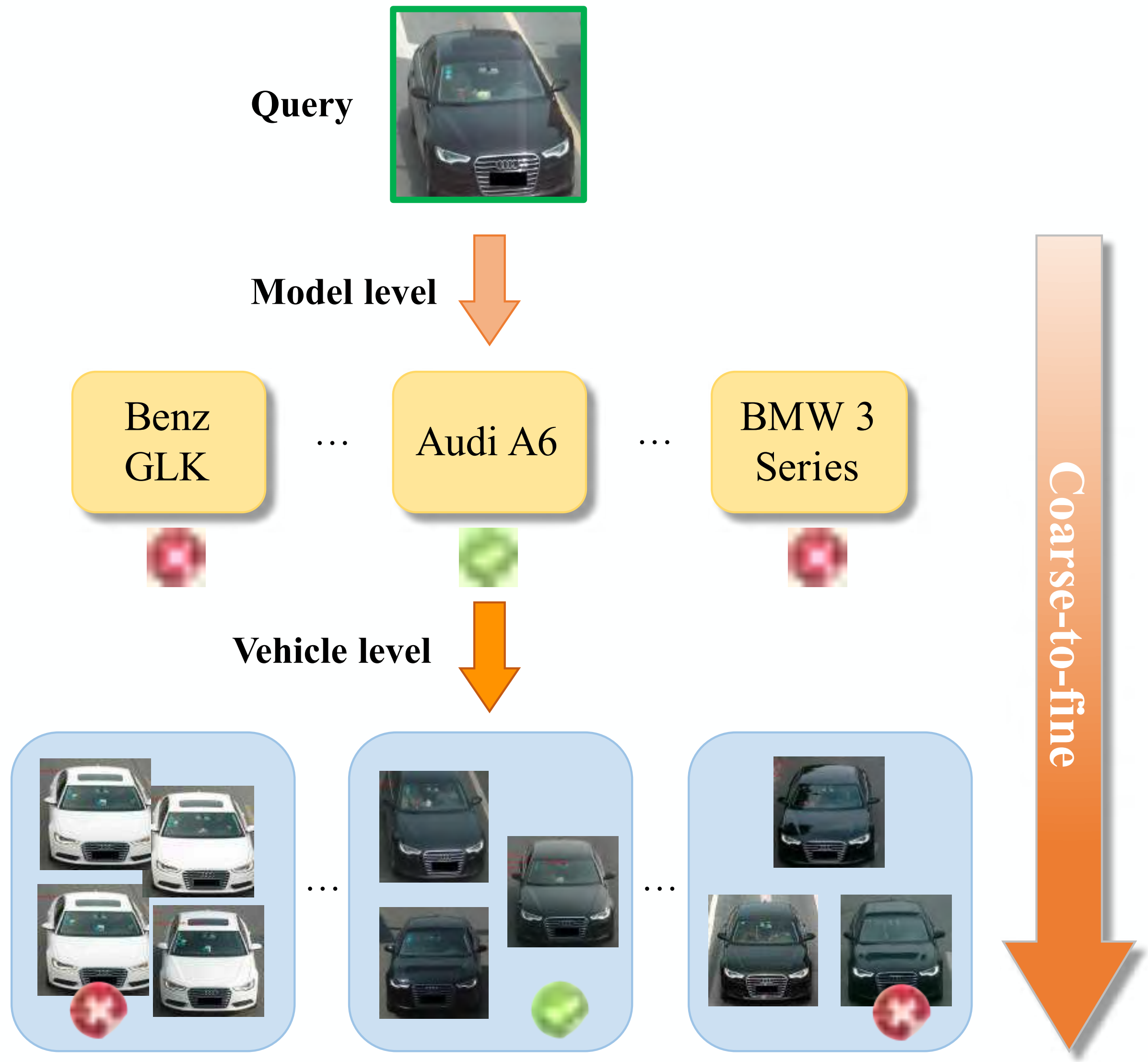
Vehicle re-identification



Other CV tasks related to fine-grained (con't)

Vehicle re-identification

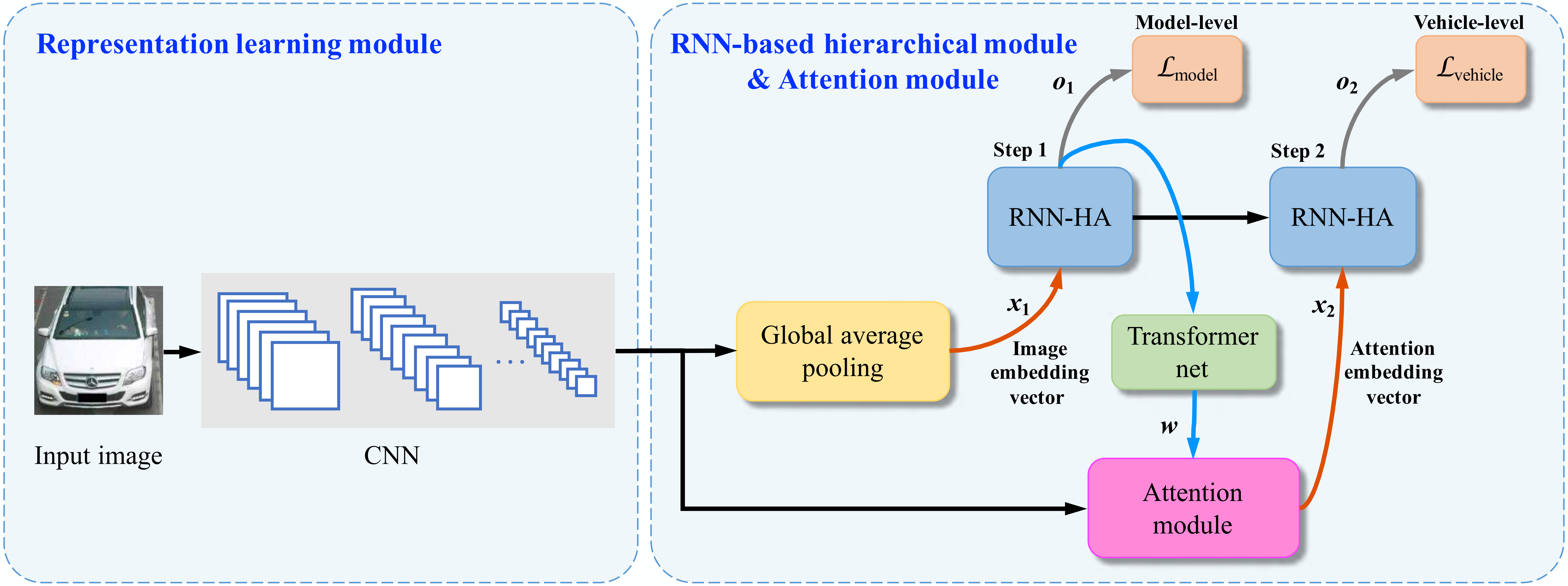
Basic idea



Other CV tasks related to fine-grained (con’t)

Vehicle re-identification

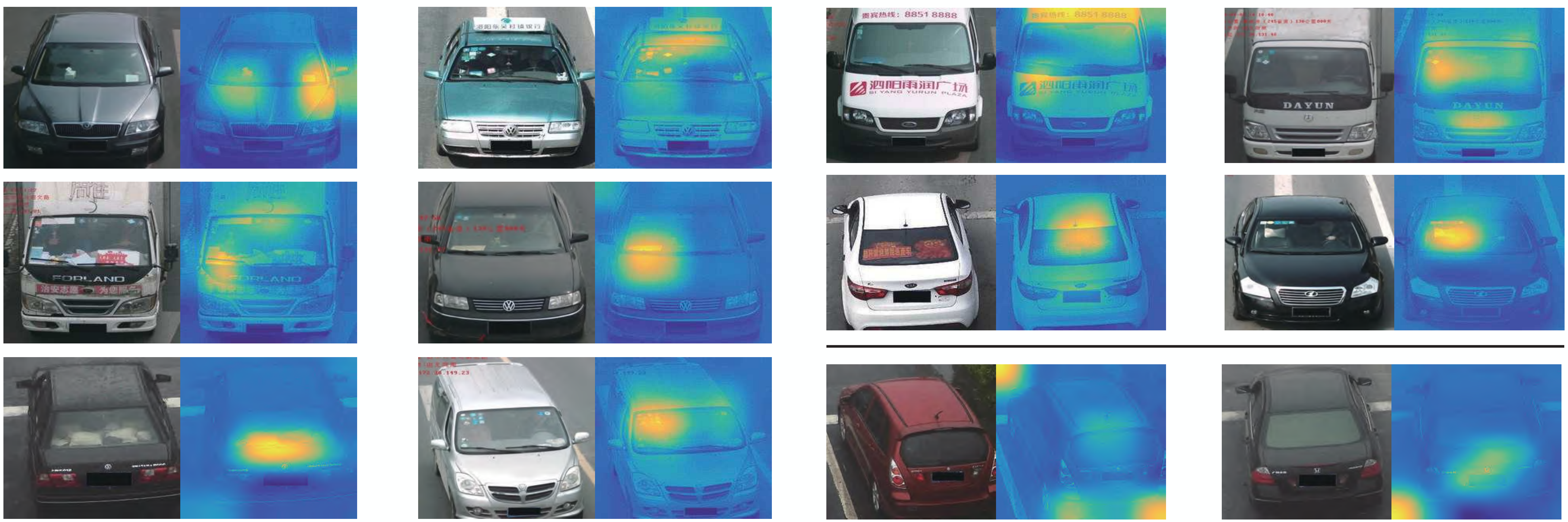
Our RNN-HA



Other CV tasks related to fine-grained (con't)

Vehicle re-identification

Qualitative results



Other CV tasks related to fine-grained (con't)

Clothes retrieval

In-shop retrieval

Consumer-to-shop retrieval

Other CV tasks related to fine-grained (con't)

Product recognition — Inventory robot

Other CV tasks related to fine-grained (con't)

Product recognition — Automatic checkout

Face++ 旷视

Other CV tasks related to fine-grained (con't)

Product recognition — Automatic Check-Out (ACO)

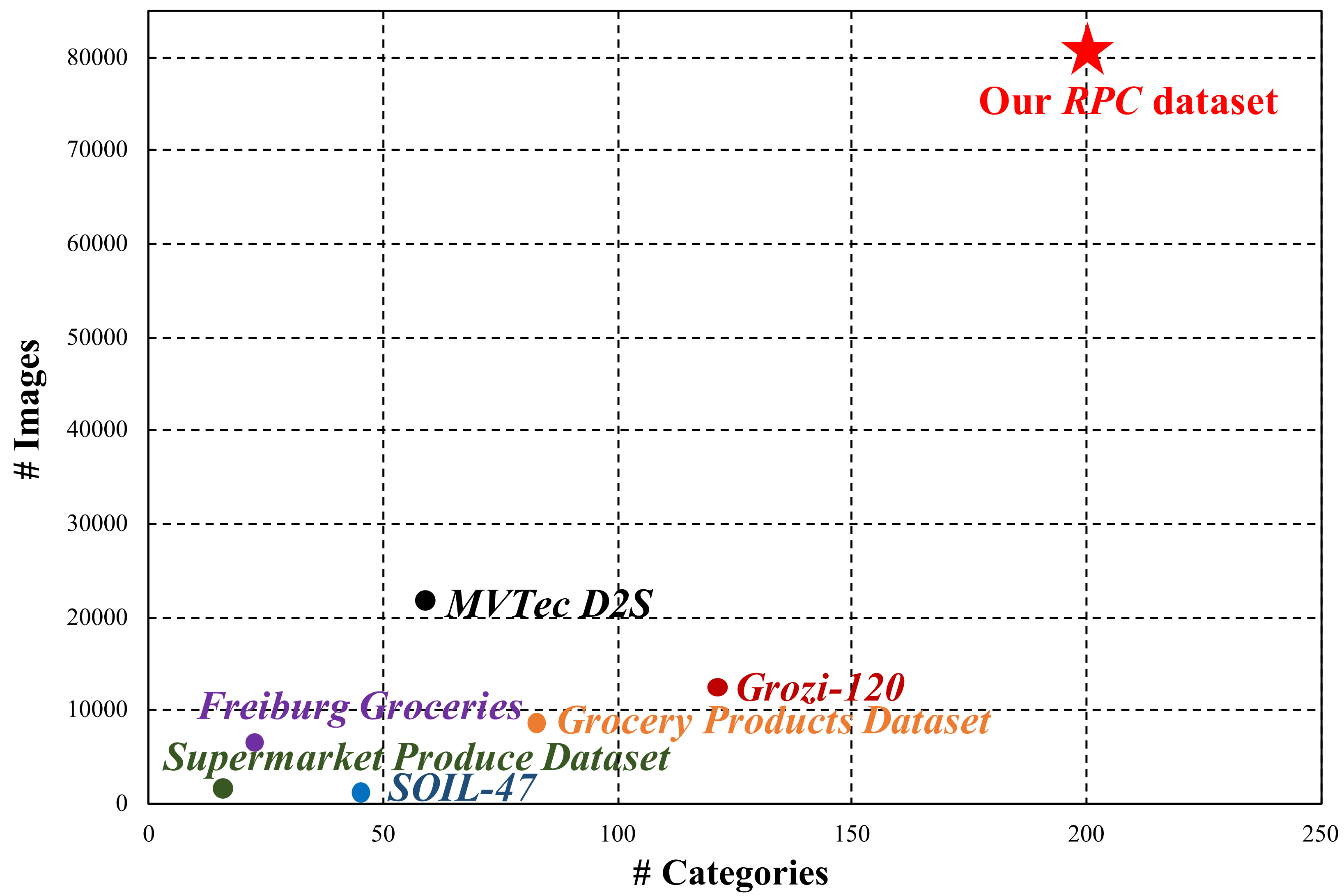
Other CV tasks related to fine-grained (con't)

Product recognition — Automatic Check-Out (ACO)

https://rpc-dataset.github.io/ http://www.weixiushen.com/

Other CV tasks related to fine-grained (con't)

Comparisons with other related datasets in the literature



Other CV tasks related to fine-grained (con't)

The images and supervisions of our task

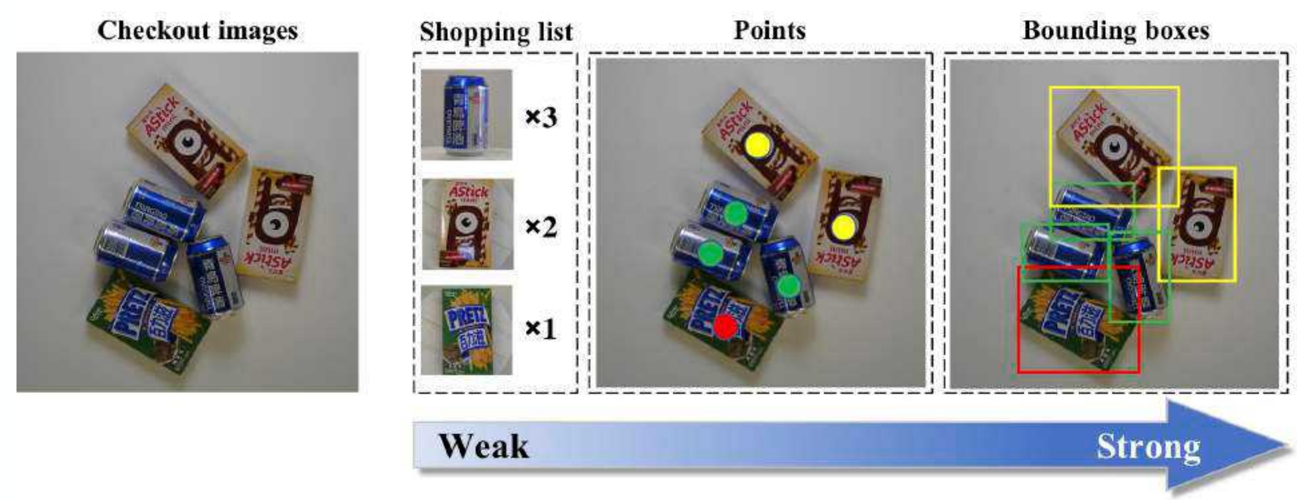
(a) Easy mode.

(b) Medium mode.

(c) Hard mode.

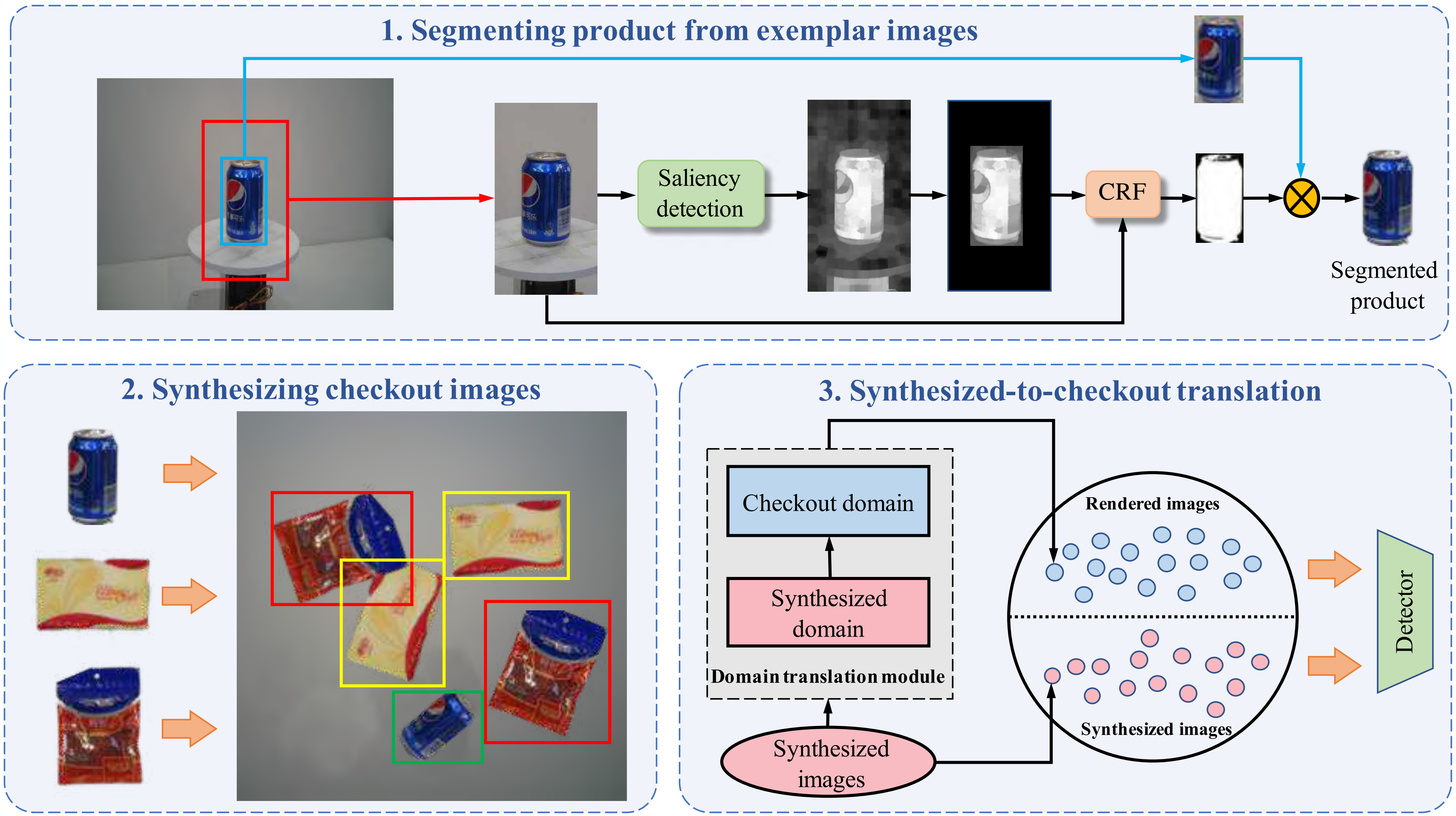
(a) Examples of bottle-like SKUs.

(b) Examples of bag-like SKUs.



Other CV tasks related to fine-grained (con't)

Our proposed baseline



Other CV tasks related to fine-grained (con't)

Main results

Table 3. Experimental results of the ACO task on our RPC dataset.

Clutter mode	*Methods*	*cAcc* (↑)	*ACD* (↓)	*mCCD* (↓)	*mCIoU* (↑)	*mAP50* (↑)	*mmAP* (↑)
Easy	Single	0.03%	8.12	1.14	2.98%	0.07%	0.01%
	Syn	18.49%	2.58	0.37	69.33%	81.51%	56.39%
	Render	63.19%	0.72	0.11	90.64%	96.21%	77.65%
	Syn+Render	**73.17%**	**0.49**	**0.07**	**93.66%**	**97.34%**	**79.01%**
Medium	Single	0.00%	16.10	1.33	1.93%	0.05%	0.01%
	Syn	6.54%	4.33	0.37	68.61%	79.72%	51.75%
	Render	43.02%	1.24	0.11	90.64%	95.83%	72.53%
	Syn+Render	**54.69%**	**0.90**	**0.08**	**92.95%**	**96.56%**	**73.24%**
Hard	Single	0.00%	20.05	1.18	0.66%	0.05%	0.01%
	Syn	2.91%	5.94	0.34	70.25%	80.98%	53.11%
	Render	31.01%	1.77	0.10	90.41%	95.18%	71.56%
	Syn+Render	**42.48%**	**1.28**	**0.07**	**93.06%**	**96.45%**	**72.72%**
Averaged	Single	0.01%	13.10	1.09	1.20%	0.06%	0.01%
	Syn	9.27%	4.27	0.35	69.65%	80.66%	53.08%
	Render	45.60%	1.25	0.10	90.58%	95.50%	72.76%
	Syn+Render	**56.68%**	**0.89**	**0.07**	**93.19%**	**96.57%**	**73.83%**

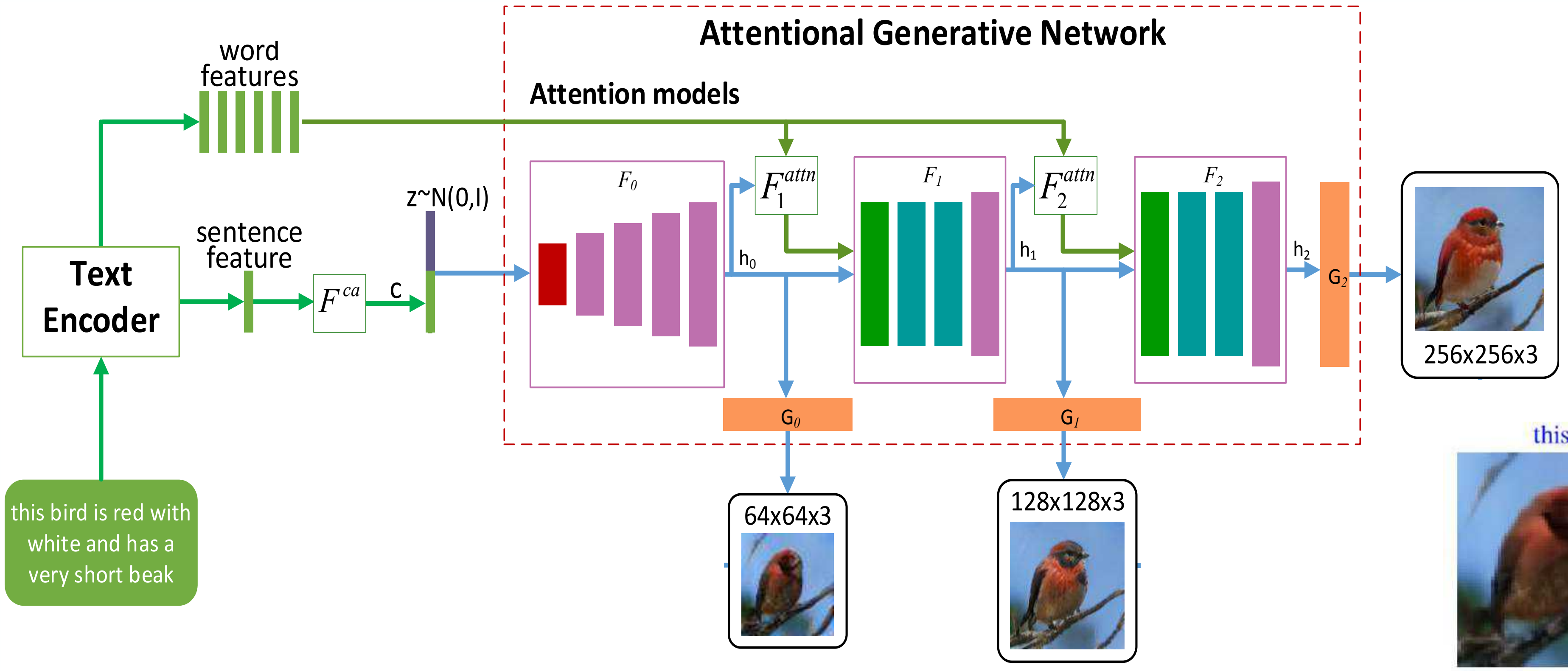
Other CV tasks related to fine-grained (con't)

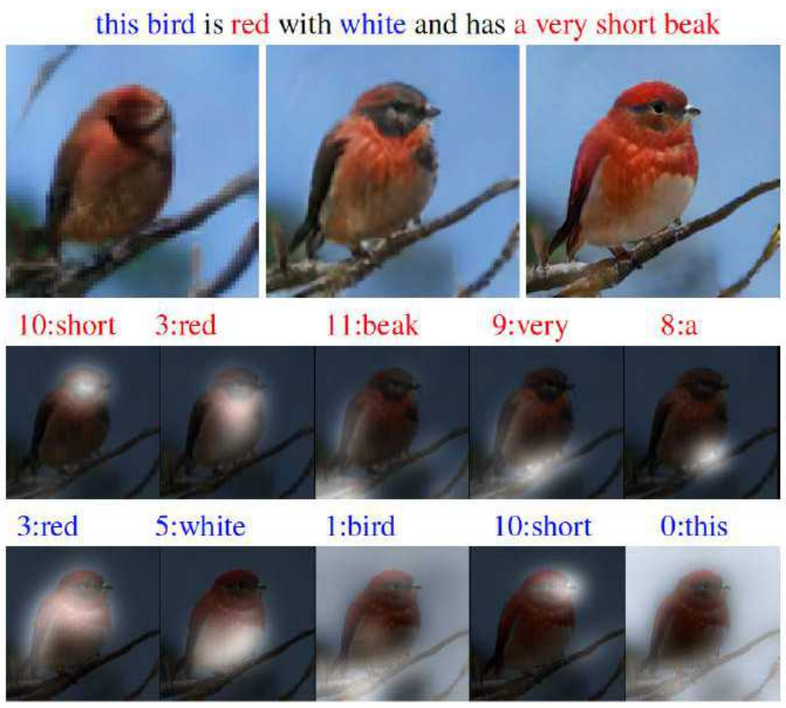
Possible research directions on our dataset

- ☑ Online learning for the ACO problem
- ☑ Multi-category object counting (with limited training samples)
- ☑ Using mixed supervision from the checkout images
- ☑ Few-shot / weakly-supervised object detection
- ☑ And many more …

New developments of fine-grained

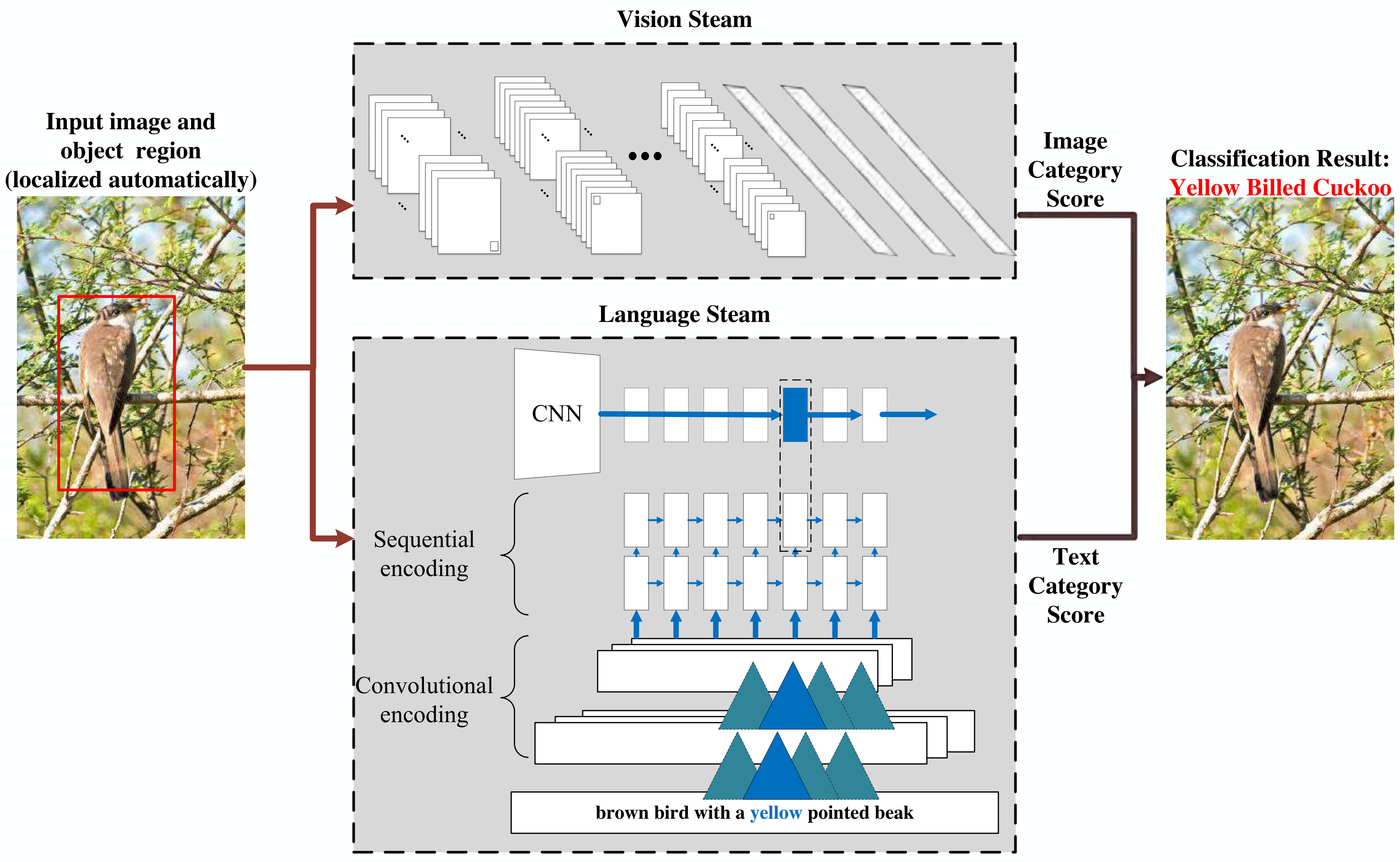
Fine-grained images with languages





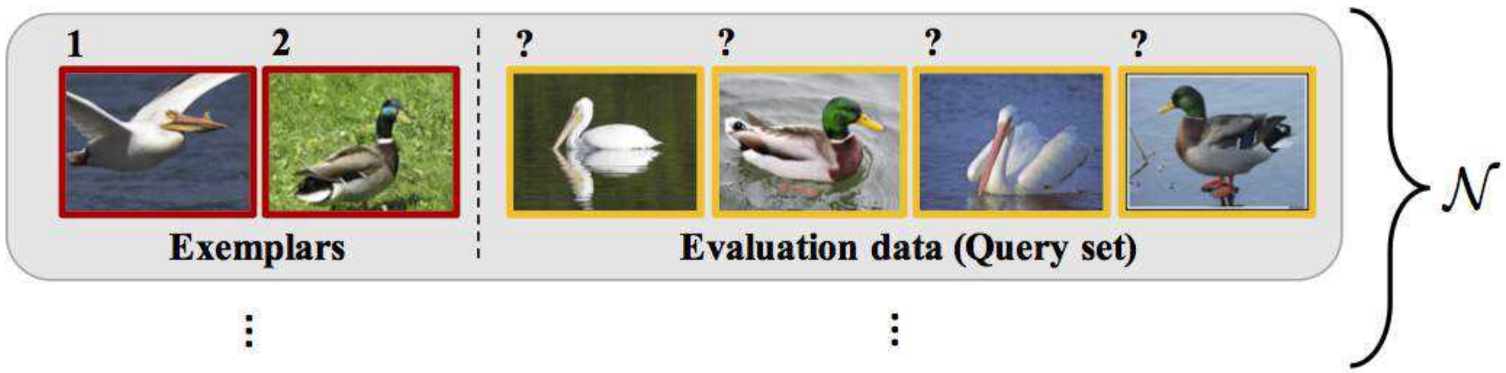
New developments of fine-grained (con’t)

Fine-grained images with languages



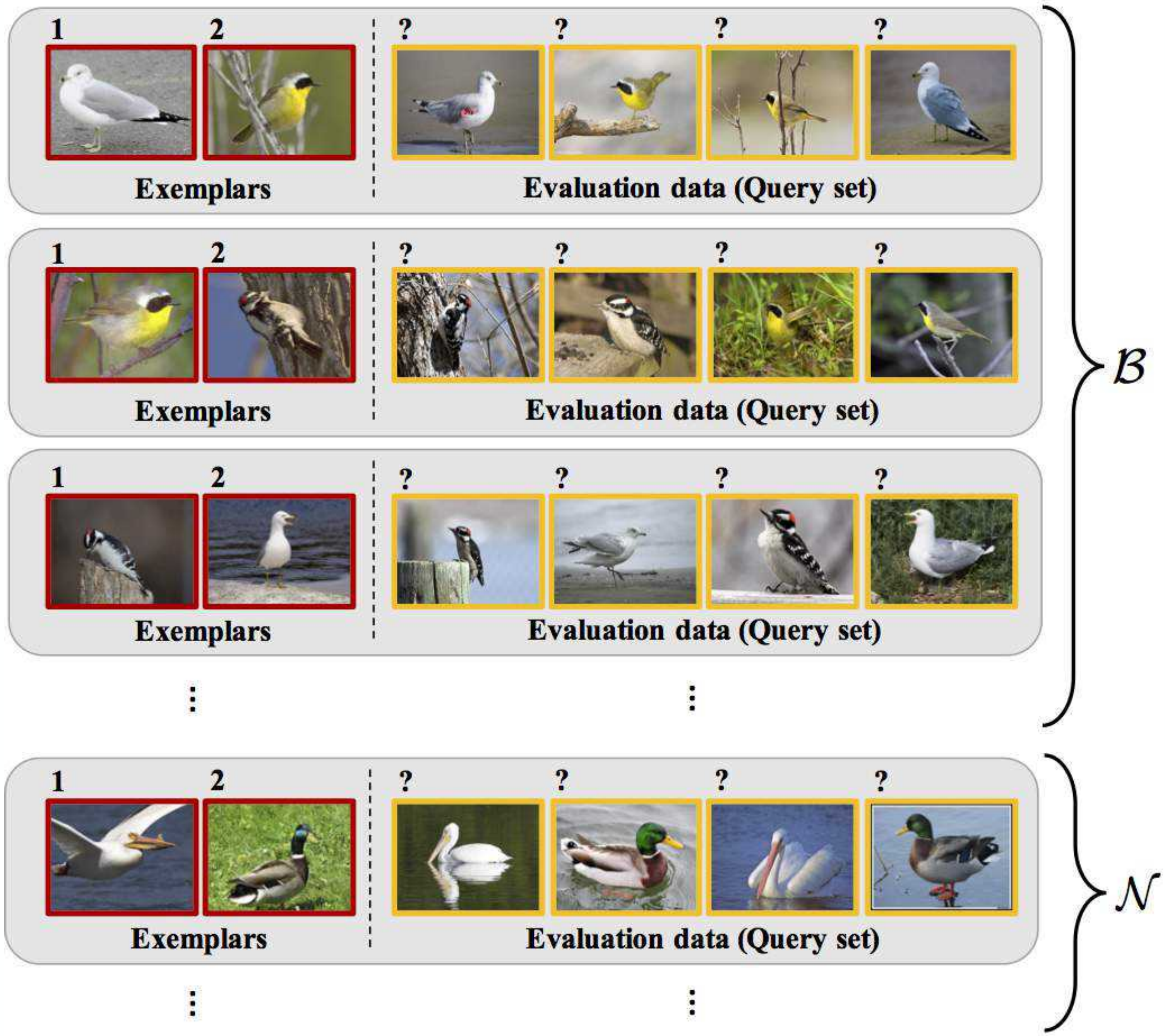
New developments of fine-grained (con't)

Few-shot fine-grained (FSFG) image recognition



New developments of fine-grained (con't)

Illustration of FSFG

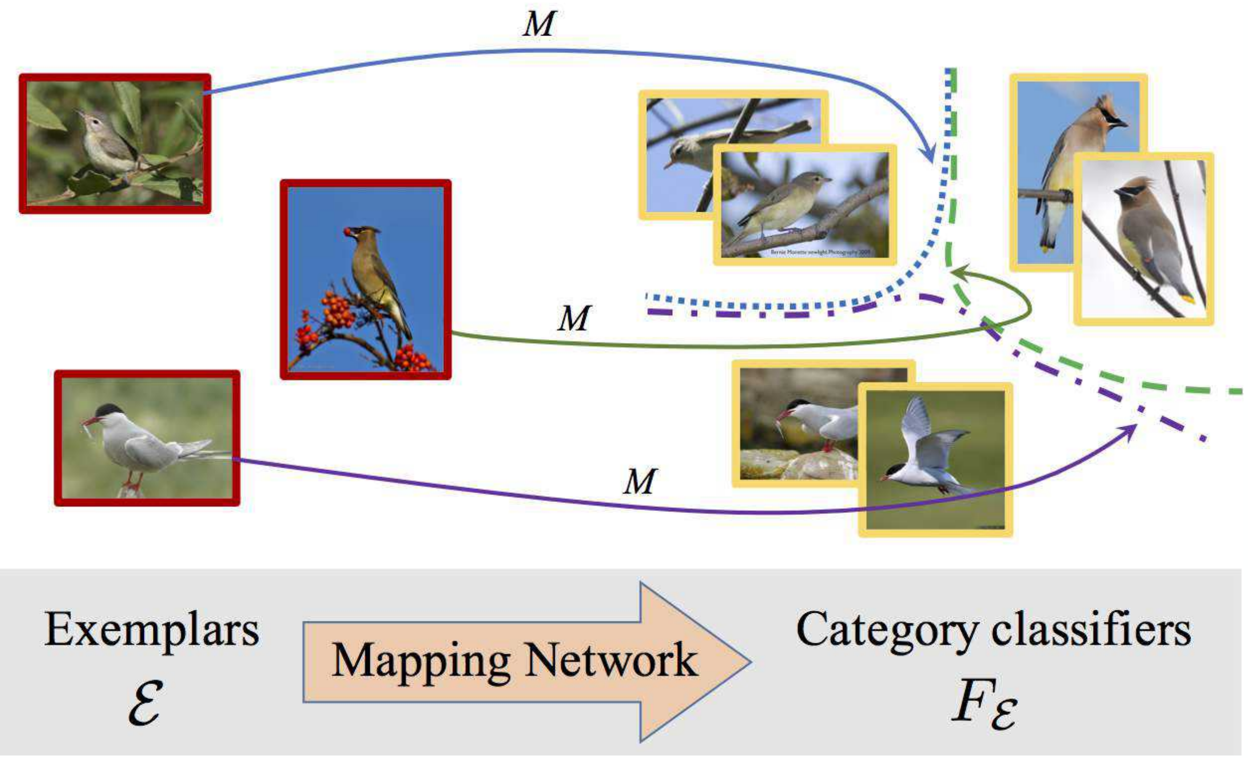


New developments of fine-grained (con't)

Learning strategy

A exemplar-to-classifier mapping function is required:

$$\mathcal{E} \xrightarrow{M} F_{\mathcal{E}} \,.$$

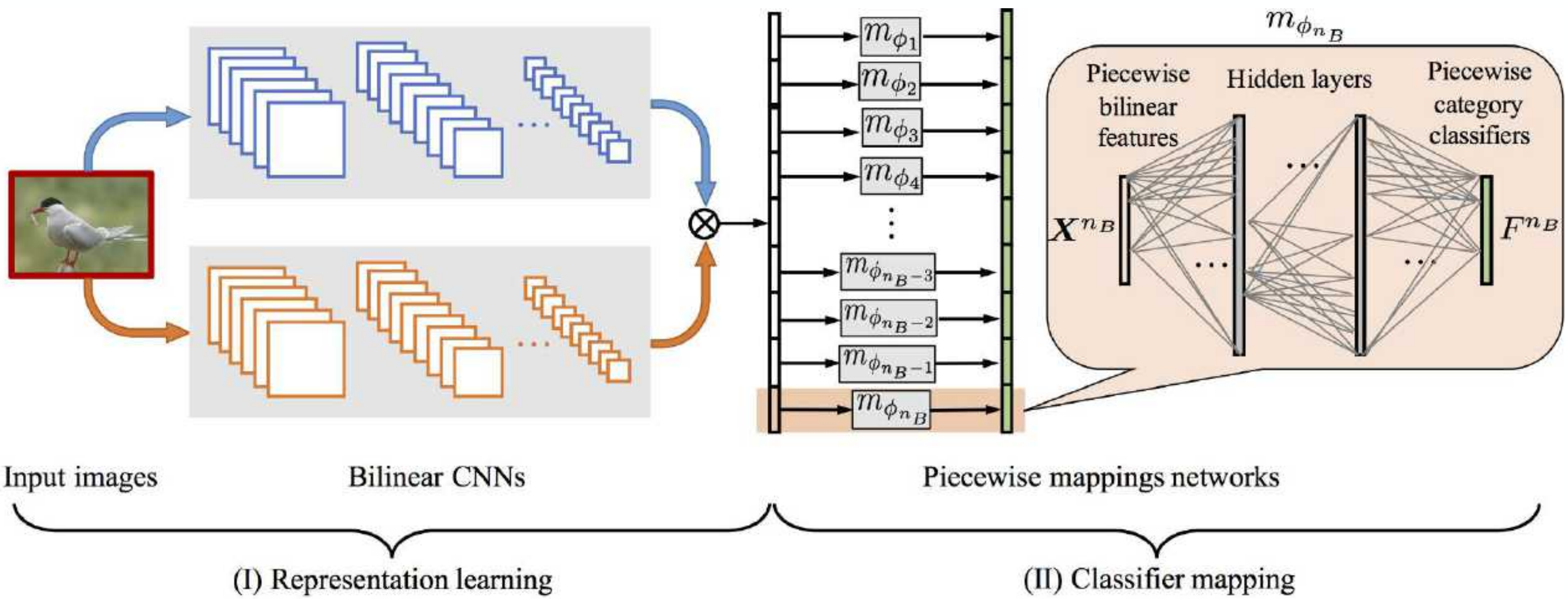


The training objective function:

$$\min_{\lambda} \underset{\{\mathcal{E},\mathcal{Q}\}\sim\mathcal{B}}{E} \{\mathcal{L}\left(F_{\mathcal{E}} \circ \mathcal{Q}\right)\}$$

New developments of fine-grained (con't)

Overview structure of our FSFG model



[Wei et al., IEEE TIP, 2019]

References

- X.-S. Wei, J.-H. Luo, J. Wu, and Z.-H. Zhou. *Selective Convolutional Descriptor Aggregation for Fine-Grained Image Retrieval*. IEEE TIP, 2017, 26(6): 2868-2881.
- C. Wah, S. Branson, P. Welinder, P. Perona, and S. Belongie. *The caltech-UCSD birds-200-2011 dataset*. Tech. Report CNS-TR-2011-001 (2011).
- J. Krause, M. Stark, J. Deng, and L. Fei-Fei. *3D object representations for fine-grained categorization*. In ICCV Workshop, Sydney, Australia, Dec. 2013.
- X.-S. Wei, C.-L. Zhang, Y. Li, C.-W. Xie, J. Wu, C. Shen, and Z.-H. Zhou. *Deep Descriptor Transforming for Image Co-Localization*. In IJCAI, Melbourne, Australia, 2017, pp. 3048-3054.
- N. Zhang, J. Donahue, R. Girshick, and T. Darrell. *Part-based R-CNNs for fine-grained category detection*. In ECCV, Part I, LNCS 8689, Springer, Switzerland, Zürich, Switzerland, Sept. 2014, pp. 834–849
- M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu. *Spatial transformer networks*. In NIPS, Montréal, Canada, Dec. 2015, pp. 2008-2016.
- T.-Y. Lin, A. RoyChowdhury, and S. Maji. *Bilinear CNN models for fine-grained visual recognition*. In ICCV, Sandiago, Chile, Dec. 2015, pp. 1449–1457.

References

- X.-S. Wei, C.-W. Xie, J. Wu, and C. Shen. *Mask-CNN: Localizing Parts and Selecting Descriptors for Fine-Grained Bird Species Categorization*. Pattern Recognition, 2018, 76:704-714.
- T. Xu, P. Zhang, Q. Huang, H. Zhang, Z. Gan, X. Huang, and X. He. *AttnGAN: Fine-Grained Text to Image Generation With Attentional Generative Adversarial Networks*. In CVPR, Salt Lake City, Jun. 2018, pp. 1316-1324.
- Y. Zhang, X.-S. Wei, J. Wu, J. Cai, J. Lu, V.-A. Nguyen, and M. N. Do. *Weakly Supervised Fine-Grained Categorization with Part-Based Image Representation*. IEEE TIP, 2016, 25(4): 1713-1725.
- J. Fu, H. Zheng, and T. Mei. *Look Closer to See Better: Recurrent Attention Convolutional Neural Network for Fine-grained Image Recognition*. In CVPR, Honolulu, Jul. 2017, pp. 4438-4446.
- H. Zheng, J. Fu, T. Mei, and J. Luo. *Learning Multi-Attention Convolutional Neural Network for Fine-Grained Image Recognition*. In ICCV, Venice, Italy, Oct. 2017, pp. 5209-5217.
- X.-S. Wei, P. Wang, L. Liu, C. Shen, and J. Wu. *Piecewise Classifier Mappings: Learning Fine-Grained Learners for Novel Categories with Few Examples*. IEEE TIP, in press.
- M. Simon and E. Rodner. *Neural activation constellations: Unsupervised part model discovery with convolutional networks*. In ICCV, pages 1143–1151, 2015.

Acknowledgment

Co-authors

Prof.
Zhi-Hua ZHOU

Prof.
Jianxin WU

Prof.
Chunhua SHEN

Dr.
Peng WANG

Dr.
Lingqiao LIU

Jian-Hao LUO

Chen-Lin ZHANG

Chen-Wei XIE

Quan CUI

Lei YANG

Researchers and engineers hiring

We are hiring self-motivated interns / full-time researchers and engineers in *computer vision* and *deep learning*. If you are interested in, please directly send your CV to my email.

weixiushen@megvii.com

MEGVII 旷视
南京研究院

Thanks all!

More resources can be found via:

http://www.weixiushen.com/