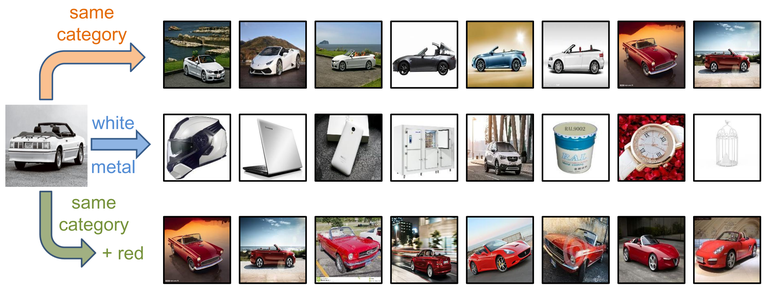
Learning Multifunctional Binary Codes for Both Category and Attribute Oriented Retrieval Tasks (DPH)

Haomiao Liu, Ruiping Wang, Shiguang Shan, Xilin Chen
Institute of Computing Technology (ICT), Chinese Academy of Sciences (CAS), China
haomiao.liu@vipl.ict.ac.cn, {wangruiping, sgshan, xlchen}@ict.ac.cn

CVPR July 21-26 HONOLULU 2017

1. Problem & Motivation

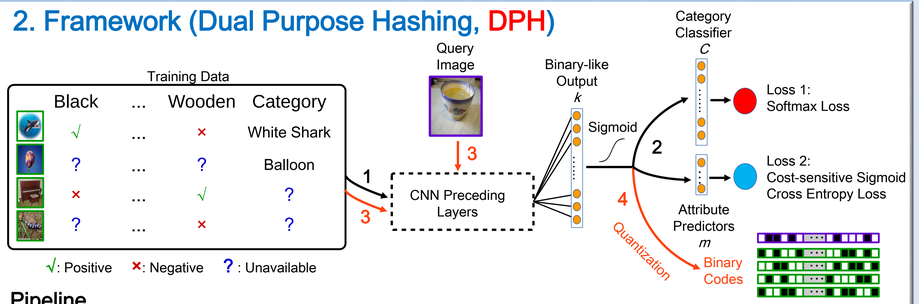
Problem: Learning unified binary codes for multiple retrieval tasks



Motivation:

- Low storage cost: images are indexed by compact binary codes
- Fast retrieval: retrieval can be done by hash table lookup or highly efficient bit operations
- Multi-functional: the learned binary codes can be used for multiple retrieval tasks

2. Framework (Dual Purpose Hashing, DPH)



Pipeline

- Train CNN models with multiple task-related loss functions and sigmoid relaxation to learn discriminative binary-like image representations.
- Quantize *k*-dim binary-like outputs to obtain binary codes.

Model description

- Compatible with most existing deep models.
- Explicitly dealing with partially labeled training data.

Source code

http://vipl.ict.ac.cn/

3. Loss Functions

y	Category label	$\mathbf{a}$	Attribute label (vector)
$\mathbf{s}$	Classifier outputs (vector)	$\mathbf{p}$	Attribute predictor outputs (vector)
C	Number of category	m	Number of attributes
$I\{\cdot\}$	Indicator function		

$$L = -\sum_{k=1}^{C} I\{y=k\}\log\left(\frac{\mathbf{s}_k}{\sum_{l=1}^{C}\mathbf{s}_l}\right)$$

softmax classification

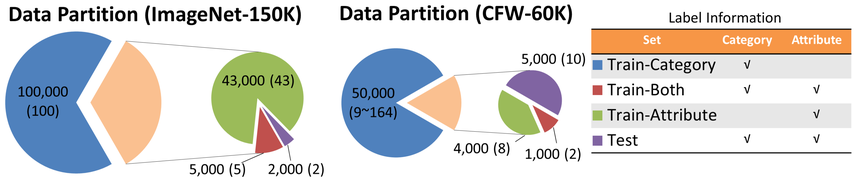
$$-\lambda\sum_{j=1}^{m} I\{\mathbf{a}_j \neq ?\}\left[w_j\mathbf{a}_j\log(\mathbf{p}_j)+(1-w_j)(1-\mathbf{a}_j)\log(1-\mathbf{p}_j)\right]$$

4. Experiments

Datasets

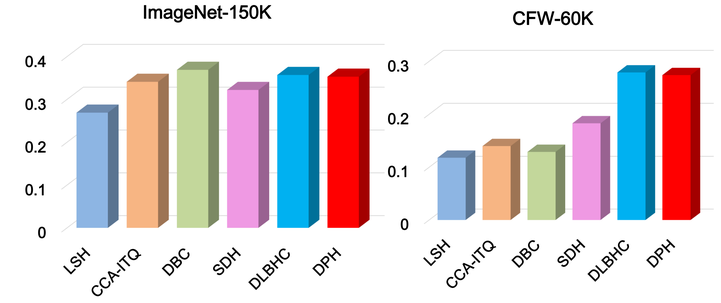
- ImageNet-150K: 15,0000 images, 50,000 have attribute labels (25 attributes).
- CFW-60K: 60,000 images, 10,000 have attribute labels (14 attributes).

Partially-labeled data (relative improvement)



Training Set	Dataset	Retrieval mAP	Attribute Prediction F1-score	Dataset	Retrieval mAP	Attribute Prediction F1-score
B	ImageNet-150K	-	-	CFW-60K	-	-
B+A		-0.009	0.103		-0.007	0.050
B+C		0.088	0.075		0.138	-0.003
B+A+C		0.095	0.126		0.146	0.060

Retrieval by Category (mAP, 256-bit binary codes)



Retrieval by Attributes (mAP, 256-bit binary codes)

	ImageNet-150K	CFW-60K
SVM-real	0.903	0.765
CNN-attribute	0.902	0.771
SVM-binary	0.871	0.729
DPH	0.868	0.804

Combined Retrieval (recall@*k*, 256-bit binary code)

5. Conclusions

This paper contributes an end-to-end hashing method for simultaneously dealing with multiple retrieval tasks

- Training the model on multiple tasks with partially labeled data can boost the performance of both tasks.
- Our method is comparable with state of the arts on individual retrieval tasks, while outperforms them on the combined task.
- Joint learning of attributes and classes is possible.