Deep Hashing with Active Pairwise Supervision

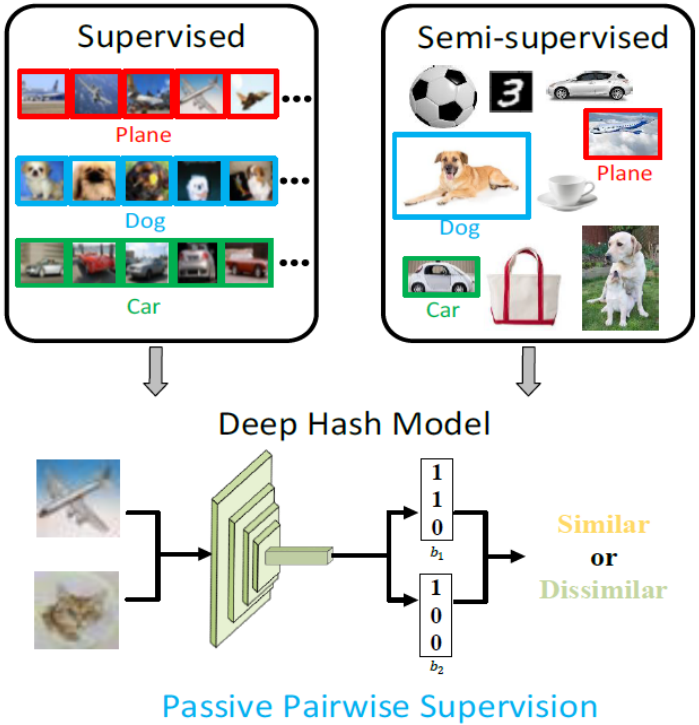
Ziwei Wang[1] Quan Zheng[1] Jiwen Lu[1] Jie Zhou[1]

[1]Department of Automation, Tsinghua University, China

i-VisionGroup@Tsinghua

Motivation

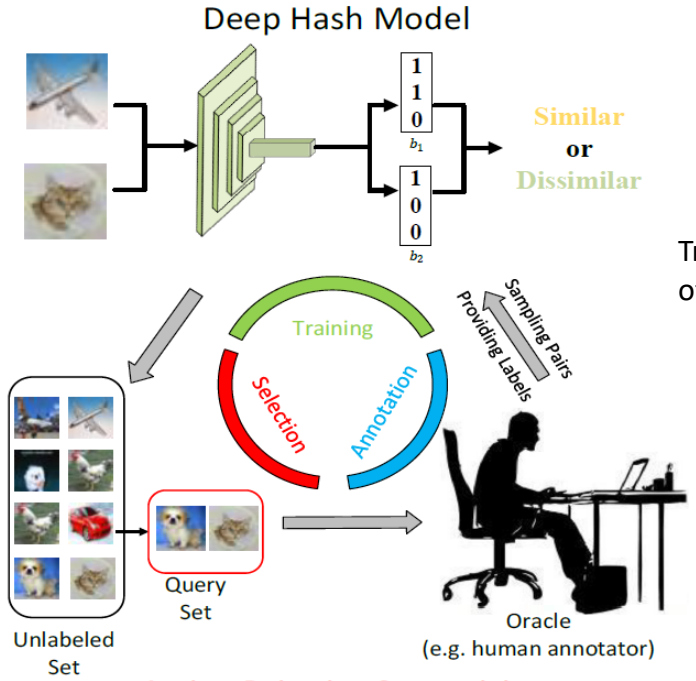
- Conventional passive deep hashing: unsupervised, supervised and semi-supervised



Methods	Annotation	Weakness
Unsupervised	None	Low discriminative power
Supervised	Full	Heavy annotation cost
Semi-supervised	Random	Ineffective supervision

Deep Hashing with Active Pairwise Supervision

- Active deep hashing: selecting samples that provide effective supervision for the Oracle to label



➢ Generalizing Structural Risk Minimization principle for sample selection:

$$\mathbb{E}(J) \leqslant \hat{\mathbb{E}}_M(J) + \Phi + MMD_S(p(\mathcal{X}), p(\mathcal{M}))$$

True risk of all data | Empirical risk of labeled data | Model complexity | Maximum mean discrepancy | Data distribution | Labeled and query data distribution

➢ Optimizing the upper bound of Structure Risk:

$$\min_{\mathcal{Q},\mathcal{H}} \quad \frac{1}{l+q} \sum_{\boldsymbol{x} \in \mathcal{L} \cup \mathcal{Q}} J + \lambda ||\mathcal{H}||_F^2 + MMD_S[p(\mathcal{X}), p(\mathcal{L} \cup \mathcal{Q})]$$

Objective for deep hashing | Regularization | MMD between all data and labeled&query data

Designing the Acquisition Function

- Defining the distance measure between hash codes of sample pairs:

$$d(\mathcal{H}(\boldsymbol{x}), \mathcal{H}(\boldsymbol{t})) = \inf_{k} \ ||\mathcal{H}(\boldsymbol{x}) - \mathcal{T}_k(\mathcal{H}(\boldsymbol{t}))||_F$$
$$= \min(||\mathcal{H}(\boldsymbol{x}_a) - \mathcal{H}(\boldsymbol{t}_a)||_F + ||\mathcal{H}(\boldsymbol{x}_b) - \mathcal{H}(\boldsymbol{t}_b)||_F,$$
$$||\mathcal{H}(\boldsymbol{x}_b) - \mathcal{H}(\boldsymbol{t}_a)||_F + ||\mathcal{H}(\boldsymbol{x}_a) - \mathcal{H}(\boldsymbol{t}_b)||_F)$$

  - Subscript a and b mean different samples in one pair

- Rewriting the MMD between two sets:

$$\inf_{\boldsymbol{k}_1, \boldsymbol{k}_2} \left\| \frac{1}{l+q} \sum_{i=1}^{l+q} \mathcal{T}_{k_{1,i}}(\mathcal{H}(\boldsymbol{x}_{1,i})) - \frac{1}{u-q} \sum_{i=1}^{u-q} \mathcal{T}_{k_{2,i}}(\mathcal{H}(\boldsymbol{x}_{2,i})) \right\|_F^2$$

  - Enumerating pairs from two sets to obtain the minimum distance

- Decoupling the acquisition function:

$$Uncertainty: \quad s_1 = \frac{1}{l+q} \sup_{y} \sum \boldsymbol{J}\boldsymbol{\alpha}$$

$$Representativeness: \quad s_2 = -\frac{l+q}{n} \mathbf{1}^u \boldsymbol{K}_{UU} \boldsymbol{\alpha}$$

$$Diversity: \quad s_3 = \frac{1}{2} \boldsymbol{\alpha}^T \boldsymbol{K}_{UU} \boldsymbol{\alpha} + \frac{u-q}{n} \mathbf{1}^l \boldsymbol{K}_{LU} \boldsymbol{\alpha}$$

J: objective, α: sample selection indicator, K_{UU}: Self-correlation of unlabeled, K_{LU} :correlation between labeled pairs and the unlabeled pairs, (l, q, n): The number of labeled, query and all pairs

- Uncertainty: Hard pairs acquire label information to provide effective supervision.
- Representativeness: Query pairs are comprehensively similar to all pairs
- Diversity: Providing sufficient information without redundancy

Experimental Results

Table 1. Comparisons of performance in different code length and annotation cost

Methods	CIFAR-10				NUS-WIDE				ImageNet-100			
	12b	24b	32b	48b	12b	24b	32b	48b	12b	24b	32b	48b
Unsupervised Hashing												
DH	22.3	23.0	23.6	23.7	22.5	23.1	23.4	23.3	12.5	13.8	14.0	14.2
GraphBit	26.9	27.2	27.0	27.3	26.7	27.0	27.2	27.4	12.9	14.5	14.7	15.1
Semi-supervised Hashing												
SSH†	35.3	37.0	38.1	38.2	30.0	31.6	35.8	32.6	19.9	21.0	21.6	23.1
SSDH	80.1	81.3	81.2	81.4	77.3	77.9	77.8	77.8	–	–	–	–
PTS^{3}H	79.8	82.8	83.5	84.3	75.2	77.4	78.3	78.9	66.1	67.5	68.0	69.7
Supervised Hashing												
DSH	61.6	65.6	66.1	67.3	54.5	55.3	55.9	56.0	47.9	50.3	50.7	51.4
DPSH	71.3	72.7	74.4	75.7	79.4	82.2	83.8	85.1	–	–	–	–
SDSH	93.9	93.9	93.9	93.4	–	81.7	82.1	82.1	–	–	–	–
Active Hashing												
DH-APS (1%)	30.5	31.9	32.6	32.8	30.1	30.6	31.2	31.8	17.9	18.1	19.5	19.6
DH-APS (⋆)	44.9	46.4	47.8	47.7	36.0	36.8	38.5	38.8	24.9	25.1	26.3	26.8
DH-APS (10%)	47.2	48.6	49.5	49.7	38.1	39.6	40.2	40.7	26.1	27.3	27.8	28.0
DH-APS (30%)	61.8	62.4	63.5	64.3	51.8	53.0	53.5	54.3	43.5	43.6	45.2	46.9
DH-APS+PTS^{3}H	82.1	85.3	86.7	86.9	79.1	81.1	82.2	82.3	68.9	70.0	70.3	71.8

Figure 1. Comparisons on precision @ distance less than 2 and top-k precision

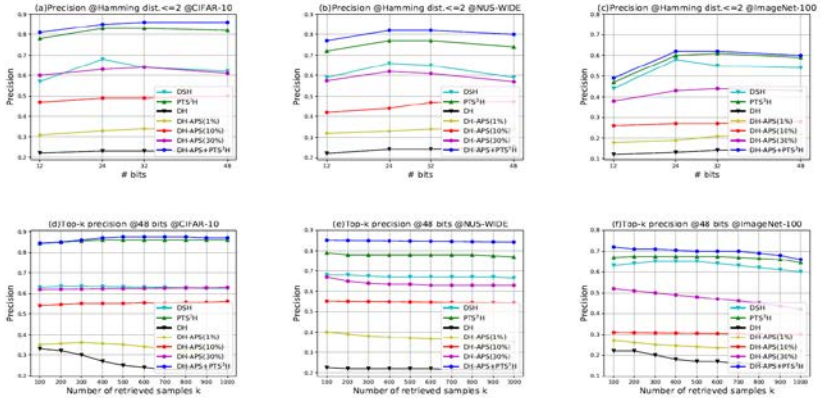


Figure 2. (a) t-SNE of sample selection and (b) visualization of chosen samples