北京航空航天大學
BEIHANG UNIVERSIYT

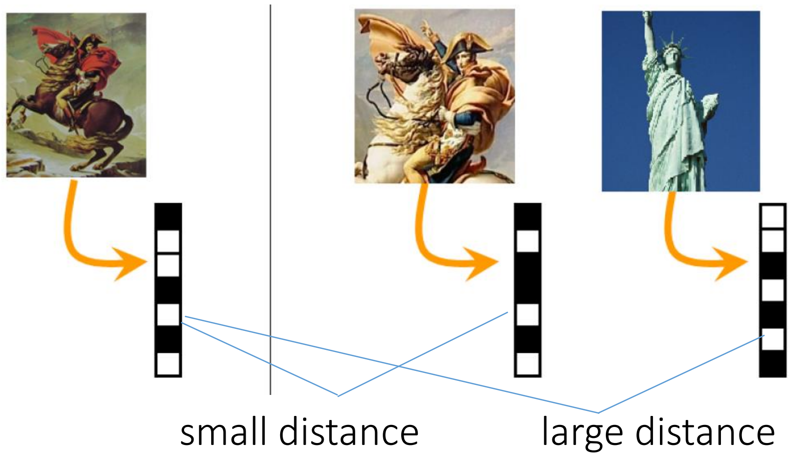
Progressive Generative Hashing for Image Retrieval

Yuqing Ma, Yue He, Fan Ding, Sheng Hu, Jun Li, Xianglong Liu

2018.7.16

CONTENT

- Hash-based Approximate Solution
  - Encode data into hash codes



  - Significantly reduce the storage, constant/sublinear query time

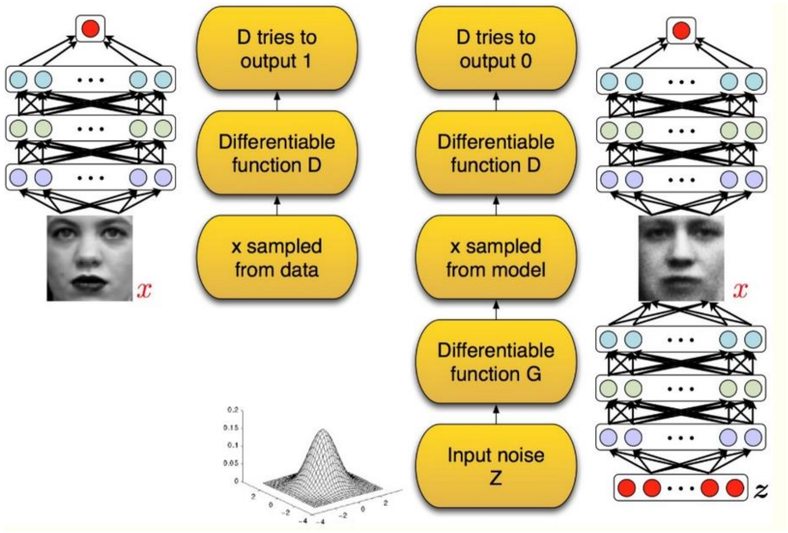
1B dataset 10K dim., using 64 bit codes

needs 8GB storage ~~40TB~~, search in Hamming distance **13s ~~15hrs~~**

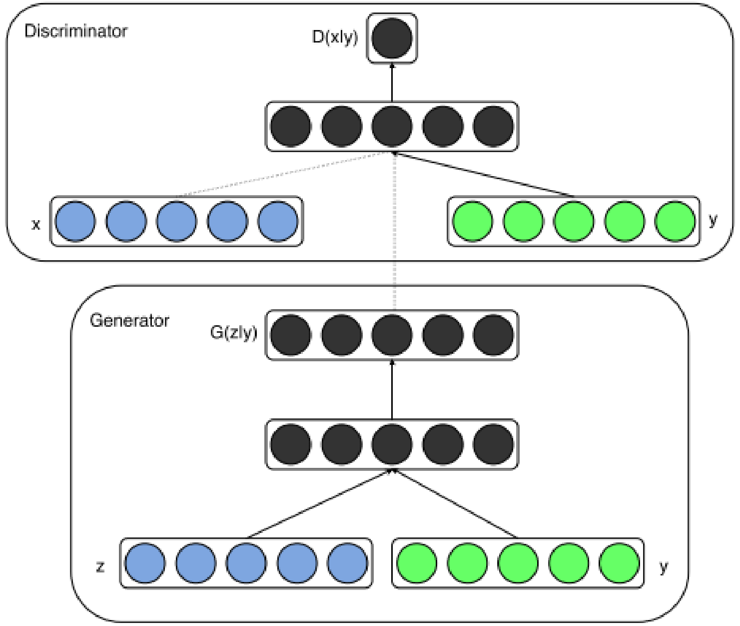
02 RELATED WORK

Generative adversarial networks

GAN



CGAN



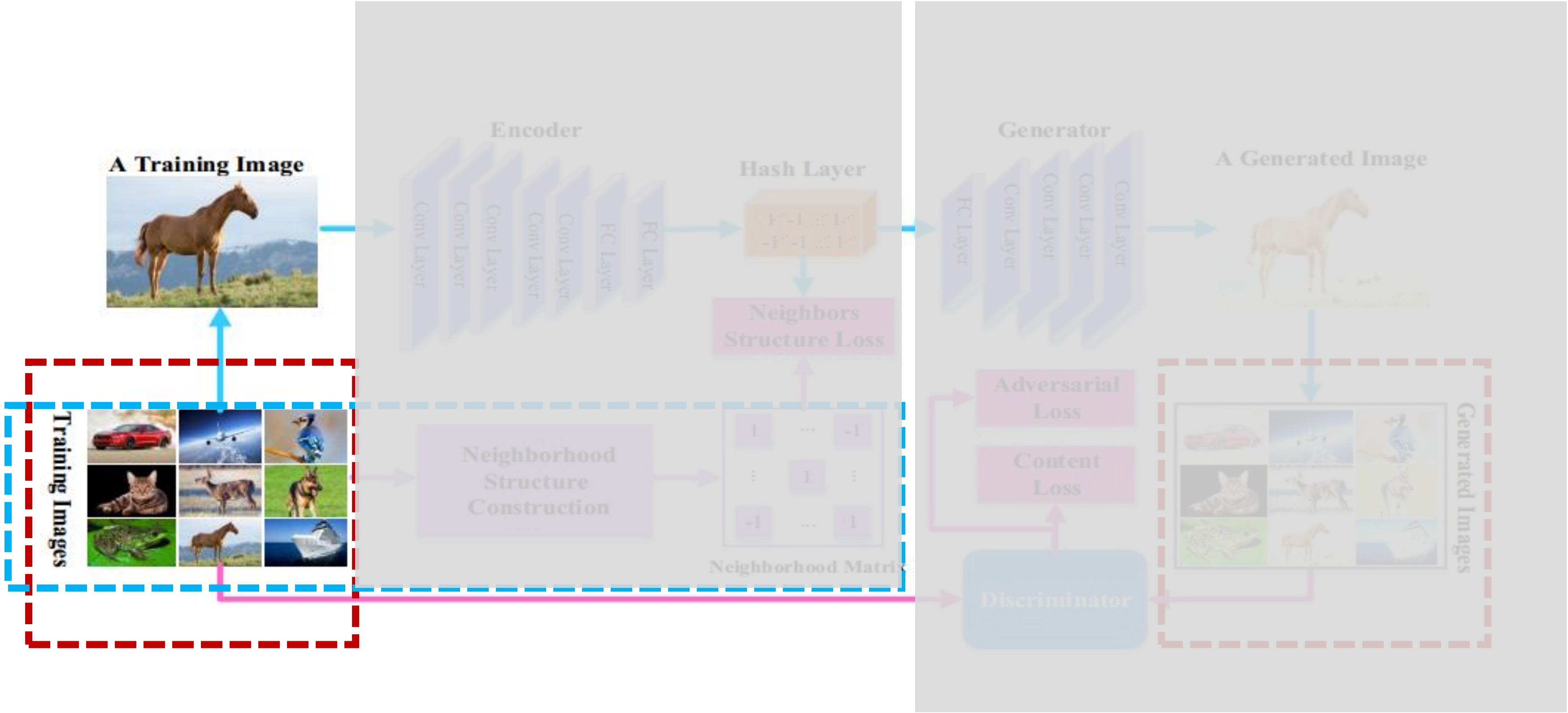
GAN: $$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))].$$

CGAN: $$\min_G \max_D V(D,G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x}|\boldsymbol{y})] + \mathbb{E}_{\boldsymbol{z}\sim p_{z}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z}|\boldsymbol{y})))].$$

02 RELATED WORK

Binary Generative Adversarial Networks for Image Retrieval

- Architecture



	mAP		
Components	24-bit	32-bit	48-bit
ℓ_N	0.487	0.511	0.543
ℓ_C	-	-	-
ℓ_A	-	-	-
$\ell_N + \ell_C$	0.247	0.379	0.497
$\ell_N + \ell_A$	0.472	0.503	0.534
$\ell_C + \ell_A$	-	-	-
$\ell_N + \ell_C + \ell_A$	**0.512**	**0.531**	**0.558**

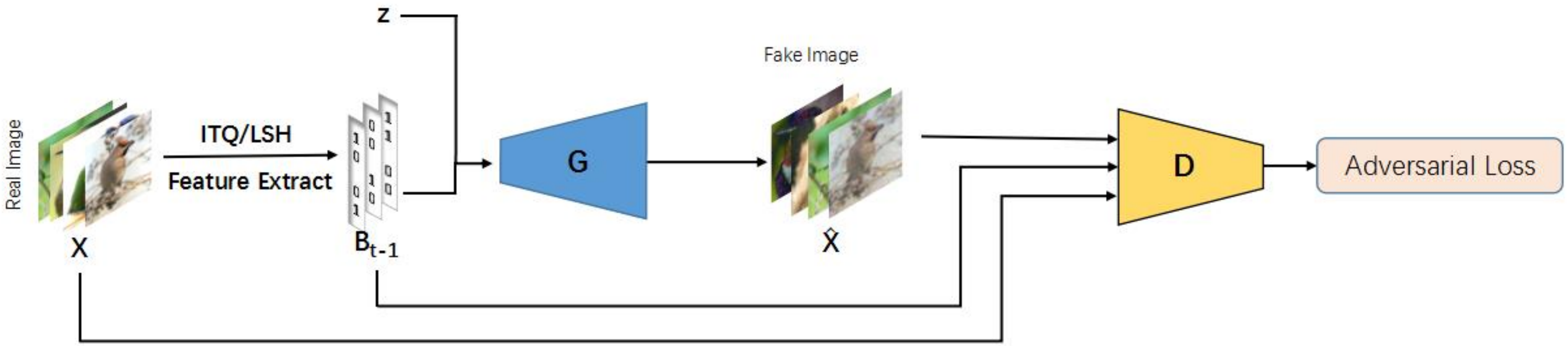
features	Resnet	GIST
mAP（%）	53.10	19.42

03 Our PGH

Hash-conditioned GANs

- Generating semantically similar images conditioned on weak binary codes

$$\min_{G}\max_{D}\ \mathcal{L}_{D,G} = \log D(\mathbf{x}_i|\mathbf{b}_i) + \log(1 - D \circ G(\mathbf{z}_i|\mathbf{b}_i))$$



03 Our PGH

Triplet Hashing

- Learning hash codes which can preserve relative similarities among the images in the triplet

$$\max\left(0, \gamma k - \|\mathbf{b}_i - \bar{\mathbf{b}}_j\|_{\mathcal{H}} + \|\mathbf{b}_i - \bar{\mathbf{b}}_i\|_{\mathcal{H}}\right),$$

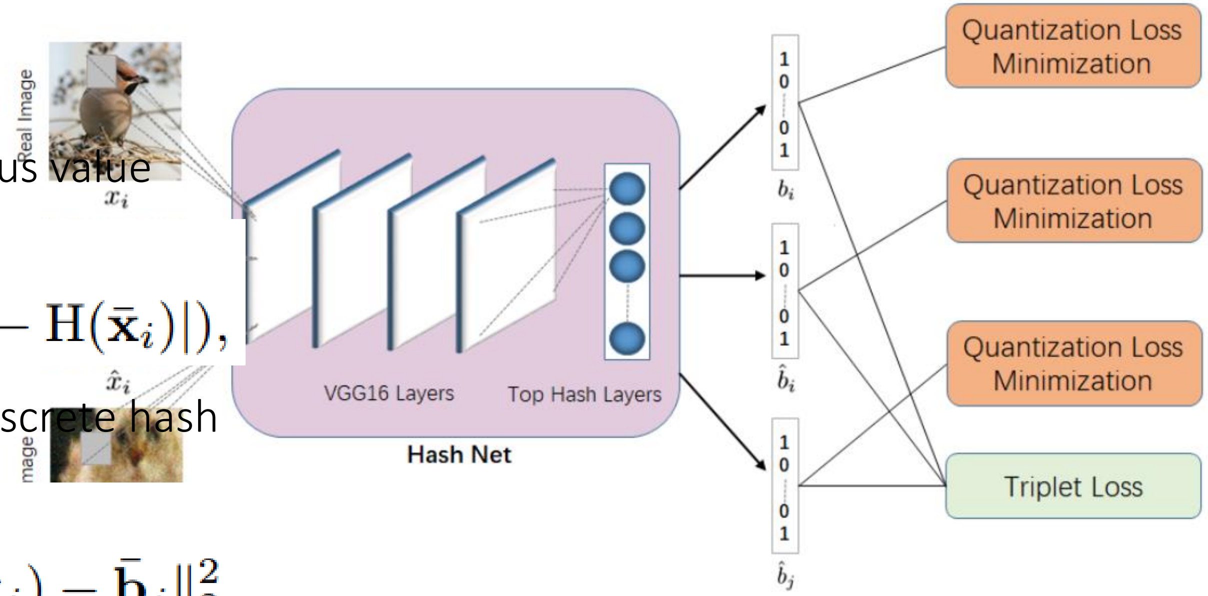
$$\mathbf{b}_i = \frac{1}{2}\mathrm{sgn}\left(\mathrm{H}(\mathbf{x}_i) - \frac{1}{2}\right) + \frac{1}{2}.$$

  - ➢ replacing the discrete binary codes with continuous value

$$h(\mathbf{x}_i, \bar{\mathbf{x}}_i, \bar{\mathbf{x}}_j) = \max(0, \gamma k - |\mathrm{H}(\mathbf{x}_i) - \mathrm{H}(\bar{\mathbf{x}}_j)| + |\mathrm{H}(\mathbf{x}_i) - \mathrm{H}(\bar{\mathbf{x}}_i)|),$$

  - ➢ Forcing the relaxation to be consistent with the discrete hash codes

$$q(\mathbf{x}_i, \bar{\mathbf{x}}_i, \bar{\mathbf{x}}_j) = \|\mathrm{H}(\mathbf{x}_i) - \mathbf{b}_i\|_2^2 + \|\mathrm{H}(\bar{\mathbf{x}}_i) - \bar{\mathbf{b}}_i\|_2^2 + \|\mathrm{H}(\bar{\mathbf{x}}_j) - \bar{\mathbf{b}}_j\|_2^2$$



- The overall loss function of deep hashing network H

$$\min_{\mathrm{H},\mathbf{B},\bar{\mathbf{B}}} \mathcal{L}_{\mathrm{H,G}} = \sum_{i,j \neq i} h(\mathbf{x}_i, \bar{\mathbf{x}}_i, \bar{\mathbf{x}}_j) + q(\mathbf{x}_i, \bar{\mathbf{x}}_i, \bar{\mathbf{x}}_j)$$

$$s.t.\ \mathbf{B}, \bar{\mathbf{B}} \in \{0,1\}^{k \times n}$$

Our PGH

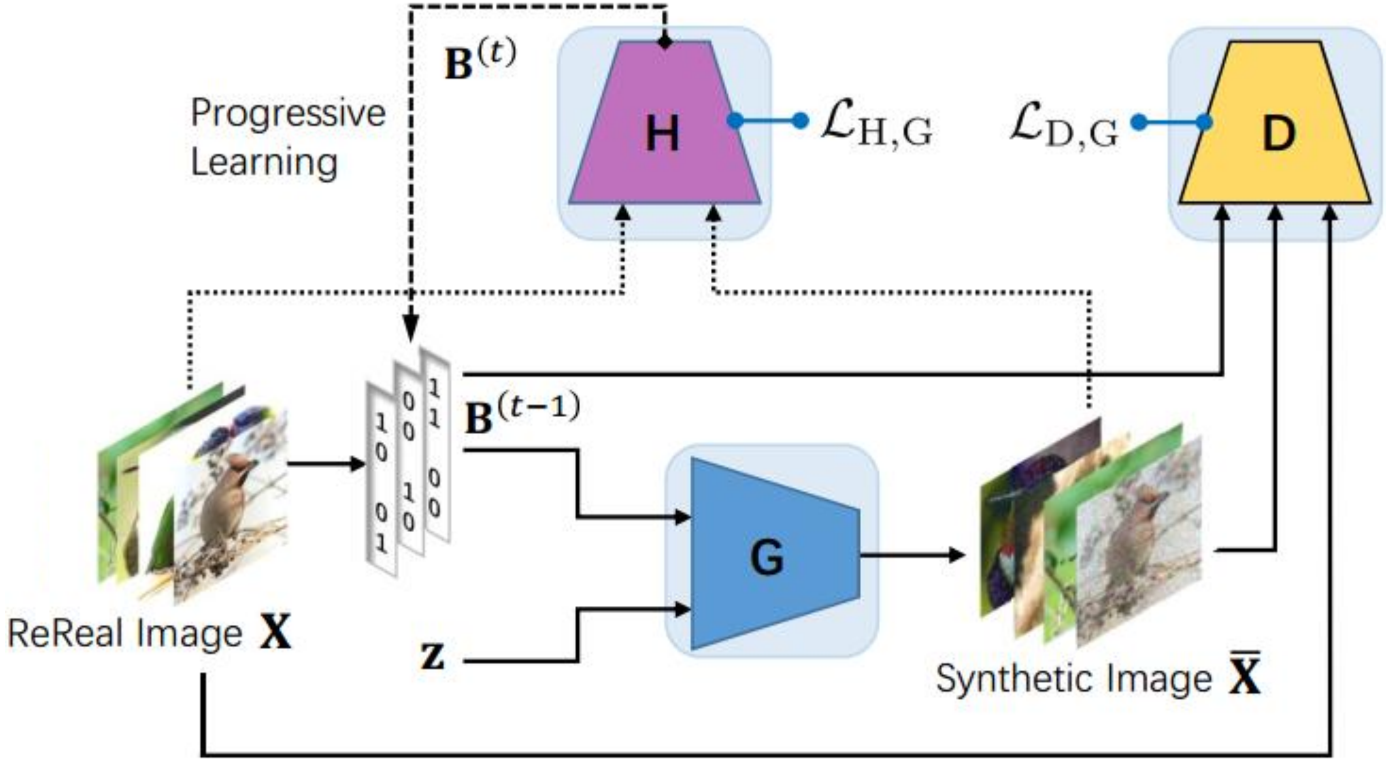
Progressive Architecture

- Feeding the learned binary codes into the hash conditioned GANs

- Progressively obtaining the improved hash codes until converging

$$\min_{\mathrm{G}^{(t)},\mathrm{H}^{(t)}} \max_{\mathrm{D}^{(t)},\mathbf{B}^{(t)},\bar{\mathbf{B}}^{(t)}} \mathcal{L}_{\mathrm{D}^{(t)},\mathrm{G}^{(t)}} + \lambda \mathcal{L}_{\mathrm{H}^{(t)},\mathrm{G}^{(t)}}$$

$$s.t. \quad \mathbf{B}, \bar{\mathbf{B}} \in \{0,1\}^{k \times n}$$



04 Experiments

Search Performance

Table 1: Performance of different methods on MNIST.

Methods	mAP (%) b = 16	mAP (%) b = 32	mAP (%) b = 64
LSH+pixel	20.52	26.02	32.14
ITQ+pixel	39.57	42.98	44.88
SH+pixel	26.60	35.92	25.01
SPH+pixel	26.49	31.03	35.59
LSH+GIST	22.75	29.08	33.95
ITQ+GIST	38.64	43.12	45.66
SH+GIST	29.76	30.55	28.50
SPH+GIST	31.15	35.57	39.27
DH+GIST	**43.14**	44.94	46.74
DeepBit	28.18	32.02	44.53
BGAN+GIST	40.26	40.78	51.61
PGH+GIST	39.20	**66.95**	**67.95**

Table 2: Performance of different methods on CIFAR-10.

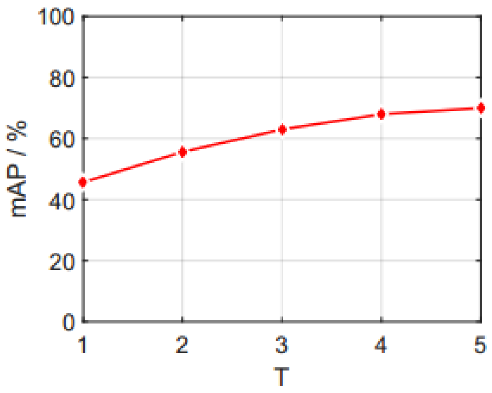
Methods	mAP (%) b = 16	mAP (%) b = 32	mAP (%) b = 64
LSH+GIST	12.78	13.97	15.07
ITQ+GIST	16.36	17.00	17.58
SH+GIST	13.06	12.92	13.07
SPH+GIST	14.63	15.14	15.86
LSH+VGG	15.38	17.17	20.73
ITQ+VGG	25.51	26.57	28.23
SH+VGG	17.29	16.44	16.56
SPH+VGG	18.97	21.26	22.97
DH+GIST	16.17	16.62	16.96
DeepBit	19.43	24.86	27.73
BGAN+GIST	19.15	19.42	21.60
PGH+GIST	**33.40**	**33.98**	**32.40**

- On MNIST:
  - shallow methods based on different features achieve comparable performance.
  - the performance of ITQ is close to deep methods
- On Cifar:
  - shallow methods with deep features significantly outperform them with hand-crafted features
  - equipped with a good feature representation, shallow methods can even beat the deep ones.

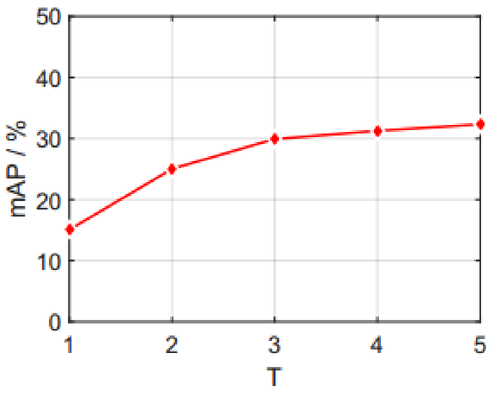
Experiments

Progressive improvement

- progressive learning curve



(a) MNIST



(b) CIFAR-10

- diffrent input hash codes

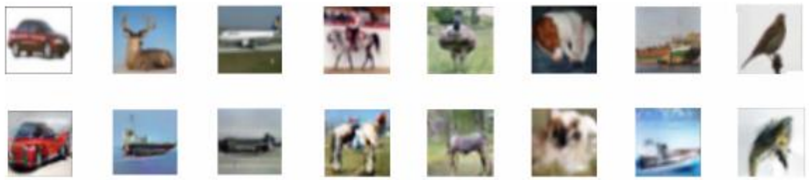
METHODS	mAP (%)				
	$t = 0$	$t = 1$	$t = 2$	$t = 3$	$t = 4$
LSH+GIST	15.07	25.03	29.92	31.26	32.18
ITQ+GIST	17.58	27.04	30.49	31.81	32.40
ITQ+VGG	28.23	32.93	33.30	-	-

04 Experiments

Semantic discovery

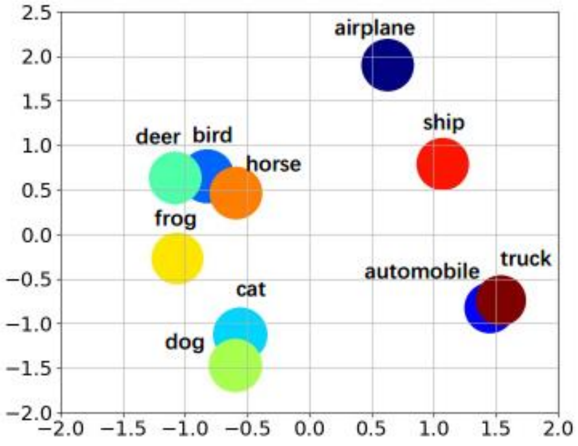
- real and synthetic images

(a) MNIST

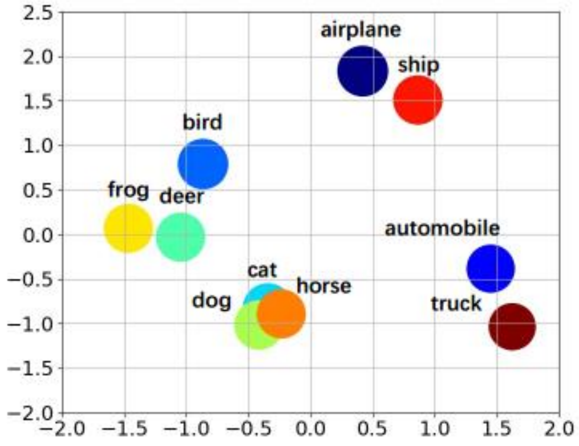


(b) CIFAR-10

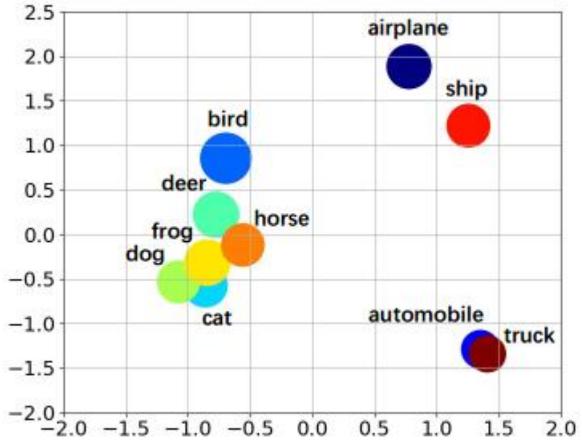
- distribution of 10 classes



(a) ITQ



(b) Deepbit



(c) PGH

Reference

1. Locality-Sensitive Hashing Scheme Based on p-Stable Distributions. Mayur Datar, Nicole Immorlica, Piotr Indyk, Vahab S. Mirrokni. *SCG*, 2004
2. Spectral Hashing. Yair Weiss, Antonio Torralba and Rob Fergus. NIPS, 2008
3. Semantic Hashing. Ruslan Salakhutdinov, Geoffrey Hinton. International Journal of Approximate Reasoning, 2009
4. Hashing with Graphs. Wei Liu, Jun Wang, Sanjiv Kumar and Shih-Fu Chang. ICML, 2010
5. Iterative Quantization: A Procerustean Approach to Learning Binary Codes. Yunchao Gong and Svetlana Lazebnik. CVPR, 2011
6. Composite Hashing with Multiple Information Sources. Dan Zhang, Fei Wang, Luo Si. ACM SIGIR, 2011
7. Generative Adversarial Networks. Ian J. Goodfellow, Jean Pougetabadie, Mehdi Mirza,Bing Xu, David Wrdefarley, Sherjil Ozair, Aaron Courville, Yoshua Bengio. 2014
8. Spherical Hashing: Binary Code Embedding with Hyperspheres. Jae Pil Heo, Junfeng He, Shih Fu Chang, Sung Eui Yoon. IEEE TPAMI, 2015
9. Conditional Generative Adversarial Nets. Mehdi Mirza, Simon Osindero. Computer Science, 2014
10. Unsupervised representation learning with Deep Convolutional Generative Adversarial Networks. Alec Radford, Luke Metz, Soumith Chintala. Computer Science, 2015
11. Deep Hashing for Compact Binary Codes Learning. Venice Erin Liong, Jiwen Lu, Gang Wang, Pierre Moulin. IEEE CVPR, 2015
12. Learning Compact Binary Desicriptors with Unsupervised Deep Neural Network. Kevin Lin, Jiwen Lu, Chu Song Chen, Jie Zhou. IEEE CVPR, 2016
13. Binary Generative Adversarial Networks for Image Retrieval. Jingkuan Song. AAAI, 2017

Q&A

Thank You!