ENHANCING FEATURE DISCRIMINATION FOR UNSUPERVISED HASHING

Tuan Hoang[+] **Thanh-Toan Do[*]** **Dang-Khoa Le Tan[+]** **Ngai-Man Cheung[+]**

[+]Singapore University of Technology and Design (SUTD) [*]University of Adelaide

Motivation

- We hypothesize that: CNNs descriptors can boost the performance of state-of-the-art hashing methods, compared to hand-crafted descriptors.
- ***"How can we further enhance discrimination of the features for hashing purpose?"***
  - Address the problem of producing compact but very discriminative features.

Contributions

- Propose to embeds feature vector into a low-dimensional vector and, simultaneously, enhances the discriminative property of features before passing them into hashing.

Gaussian Mixture Model embedding (Gemb)

1. Dimensionality reduction

- Reduce computational cost
- Produce efficient hashing codes.
  - Balance hash code.
  - Bits are pairwise independent.
- Accomplished by PCA:independent

$$\tilde{X} = \{\tilde{x}^{(1)}, \dots, \tilde{x}^{(m)}\} \in \mathbb{R}^d \rightarrow X = \{x^{(1)}, \dots, x^{(m)}\} \in \mathbb{R}^D$$

Where $D \ll d$.

Fig. 1: Our proposed **Gemb** method:

+ The **inputs** are image global descriptors such as GIST or CNN.

+ The **outputs** are the embedding features for **hashing**.

2. Posterior probability as embedding features

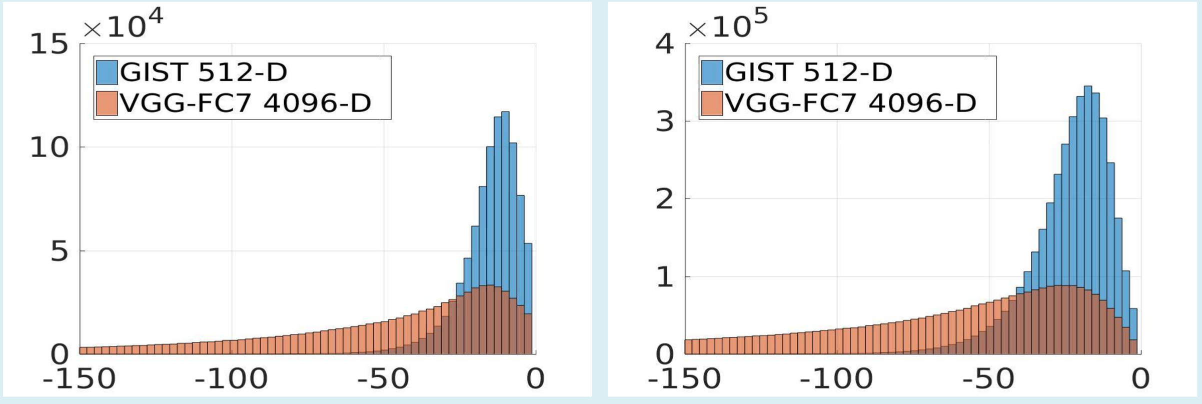
- Assume data belonging to a Gaussian Mixture Model (GMM): $\lambda = \{w_i, \mu_i, \Sigma_i\}_{i=1,\dots,N}$.
- The embedding feature of a sample $x^{(t)}$:

$$z^{(t)} = \left[P(j|x^{(t)}, \mu_j, \Sigma_j); j = 1 \dots N\right]$$

  - The posterior probability captures the strength of relationship between a sample $x^{(t)}$ and a Gaussian model $N(\mu_i, \Sigma_i)$.

$$P(j|x^{(t)}, \mu_j, \Sigma_j) = \frac{w_j p_j(x^{(t)}|\mu_j, \Sigma_j)}{\sum_{i=1}^{N} w_i p_i(x^{(t)}|\mu_i, \Sigma_i)}$$

3. *"Unsparsifying"* by Power Normalization

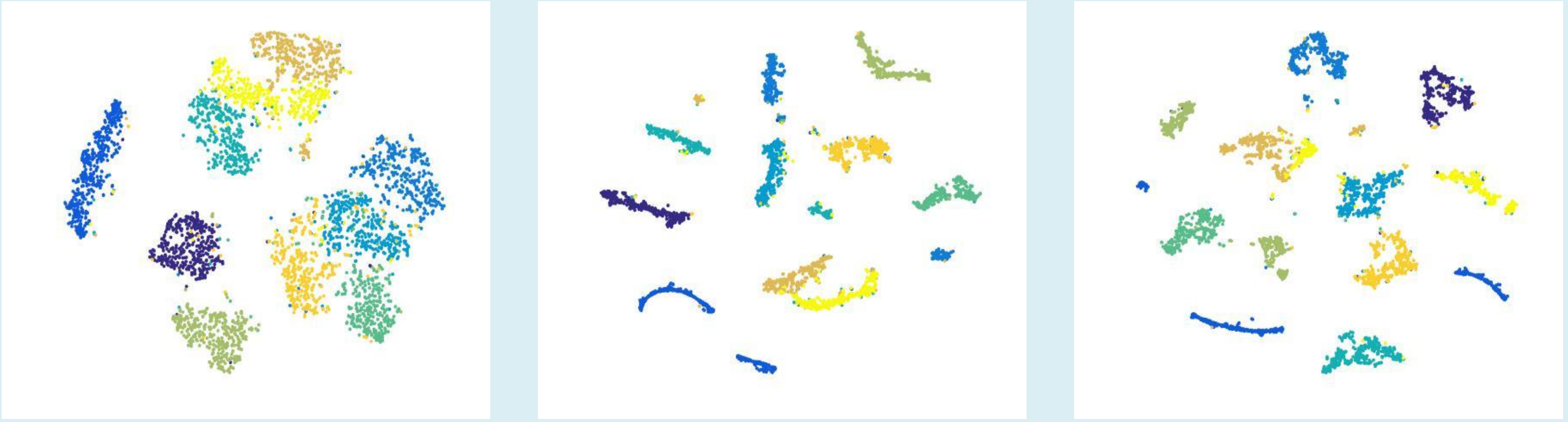


(a) N = 16 (a) N = 64

Fig 2. Histogram of embedding features (in log scale) for CIFAR-10.

- Observation:
  - Number of Gaussians ~ Embedding features sparsity.
  - Global descriptors discrimination ~ Embedding features sparsity.
- Unsparsifying: $f(x) = \mathrm{sign}(x)|x|$

Visualize descriptors



(a) Without Gemb (b) Gemb (N = 32) (c) Gemb (N = 64)

Fig 3. Visualizing GIST descriptors of a subset of MNIST dataset.

Results

	Methods	mAP					
		CIFAR-10			LabelMe-12-50k		
		16	32	64	16	32	64
GIST 512-D	SH	12.88	12.71	12.99	10.74	10.76	10.95
	SpH	14.46	15.13	15.88	11.86	13.02	13.67
	BA	15.34	16.86	17.74	14.21	14.55	15.43
	ITQ	16.59	17.42	18.02	15.07	16.06	16.58
	Gemb+BA	20.80	22.20	22.45	19.95	21.64	21.69
	Gemb+ITQ	**21.36**	**22.44**	**22.59**	**20.79**	**21.69**	**21.92**
VGG-FC7	SH	18.31	16.54	15.78	12.60	12.59	12.24
	SpH	18.82	20.93	23.40	13.59	15.10	17.03
	BA	25.38	26.16	27.99	16.96	18.42	20.80
	ITQ	26.82	27.38	28.73	18.06	19.40	20.71
	Gemb+BA	27.24	28.52	29.97	22.63	24.05	24.19
	Gemb+ITQ	**27.61**	**29.12**	**30.01**	**23.37**	**24.26**	**25.37**

Table 1. Retrieval performance (mAP).

- VGG-FC7 descriptors achieve higher hashing performances GIST 512-D descriptors.
- Gemb clearly helps to boost performance of BA and ITQ significantly.

Reference

- Link to code: https://github.com/hnanhtuan/Gemb