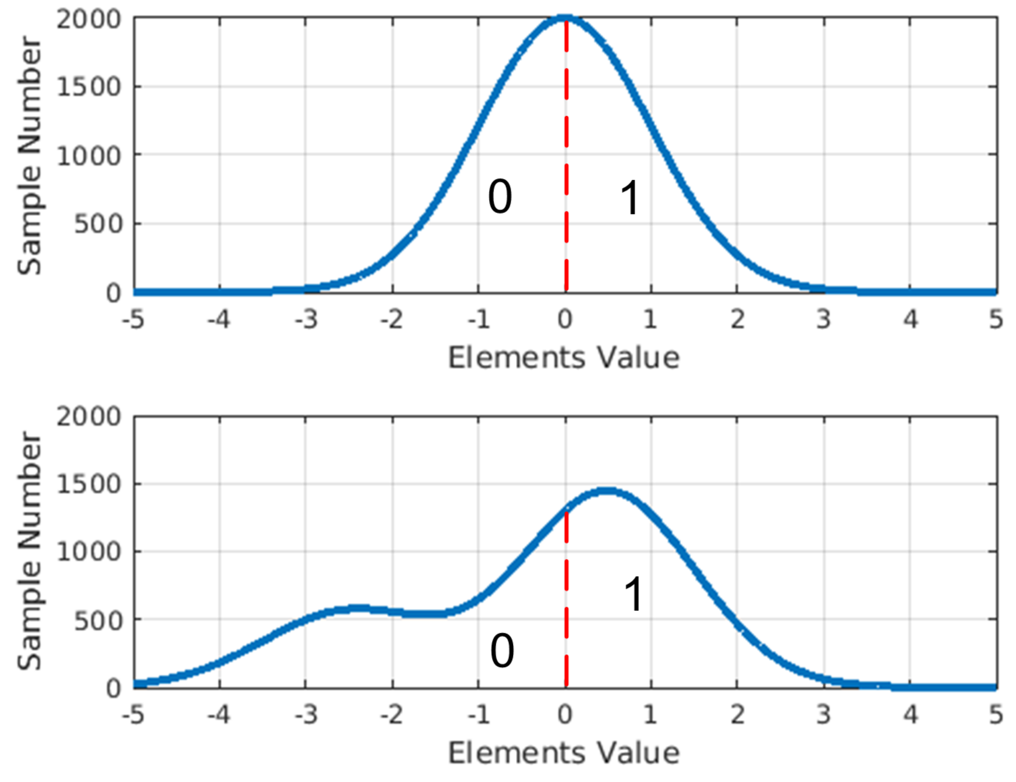
Learning Deep Binary Descriptor with Multi-Quantization

Yueqi Duan, Jiwen Lu, Ziwei Wang, Jianjiang Feng, Jie Zhou

IEEE 2017 Conference on Computer Vision and Pattern Recognition

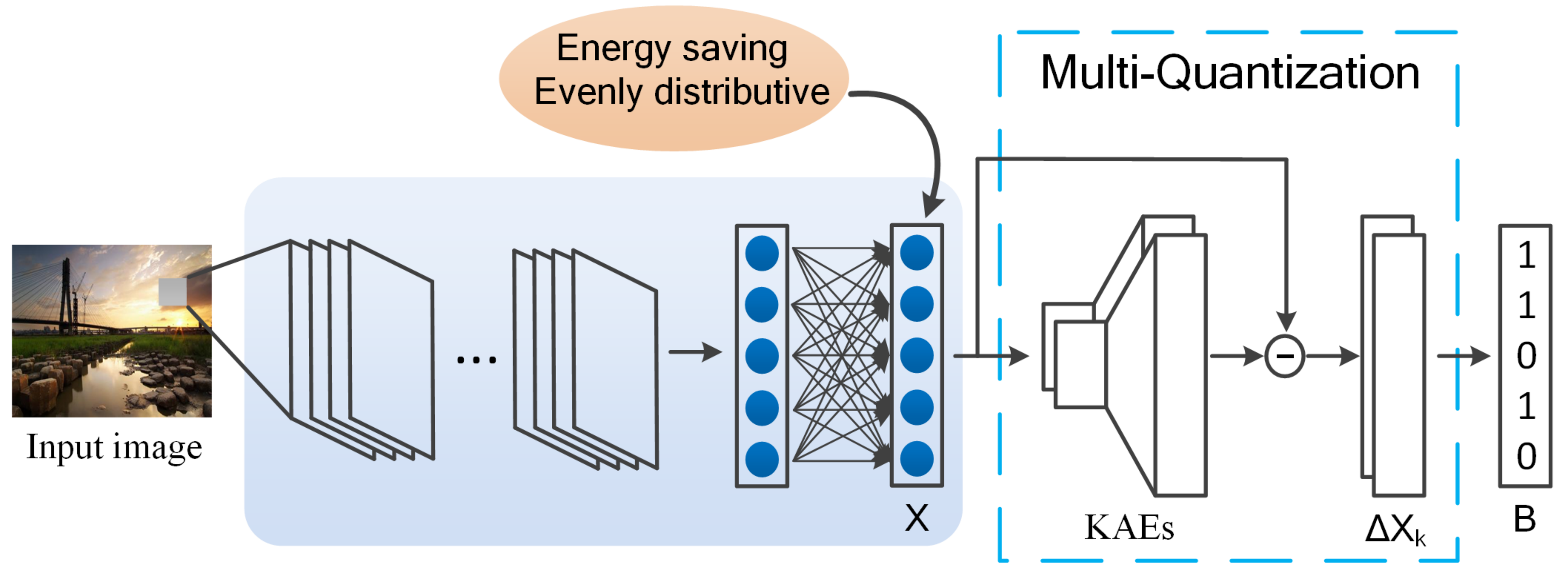
Motivation

- Deep binary descriptors show **strong discriminative power** and **low computational cost** [1]



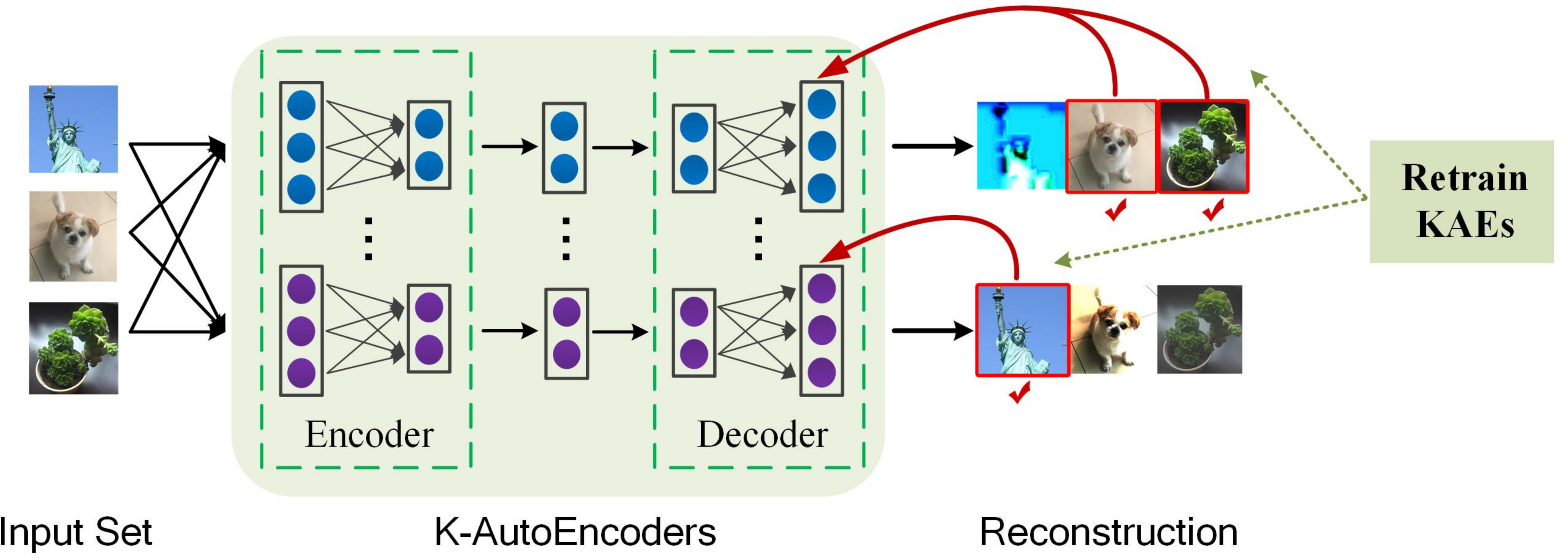
- There are two key limitations of the sign function based binarization:
  - Zero is not the proper threshold for many **data distributions** [2]
  - Existing binarization approaches perform on each bit individually, which ignore the holistic information

Flowchart



- Consider the binarization as a multi-quantization task
- Jointly learn the parameters and the binarization

K-AutoEncoders (KAEs)



- Associate each image with an autoencoder with the minimum reconstruction error
- Retrain KAEs with the corresponding images

Objective Function

$$\begin{aligned}\min_{\mathbf{X},\mathbf{W}_k} J &= J_1 + \lambda_1 J_2 + \lambda_2 J_3 \\ &= \sum_{n=1}^{N} \varepsilon_{nk_n}^2 + \lambda_1 \sum_{k=1}^{K} \sum_{l} \|\mathbf{W}_k^{(l)}\|_F^2 \\ &- \lambda_2 \mathrm{tr}((\mathbf{X}-\mathbf{U})^T(\mathbf{X}-\mathbf{U})).\end{aligned}$$

- J1: Minimize the reconstruction error of the features
- J2: Regularization term of KAEs
- J3: Enlarge the variance of each element

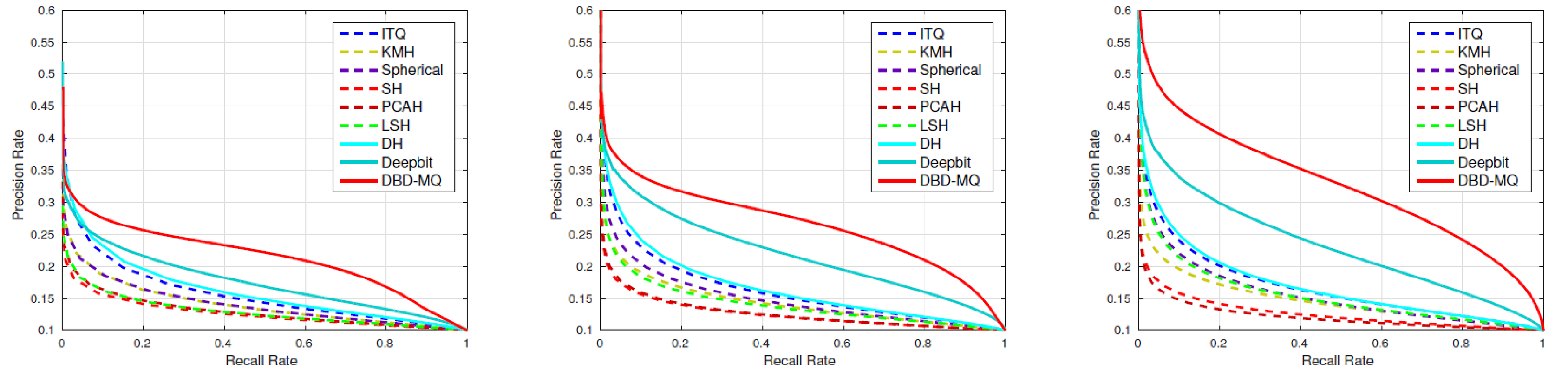
References

[1] Kevin Lin, Jiwen Lu, Chu-Song Chen, and Jie Zhou, Learning compact binary descriptors with unsupervised deep neural networks, CVPR, 2016.

[2] Jun Zhang, Jimin Liang, and Heng Zhao, Local energy pattern for texture classification using self-adaptive quantization thresholds, TIP, 2013.

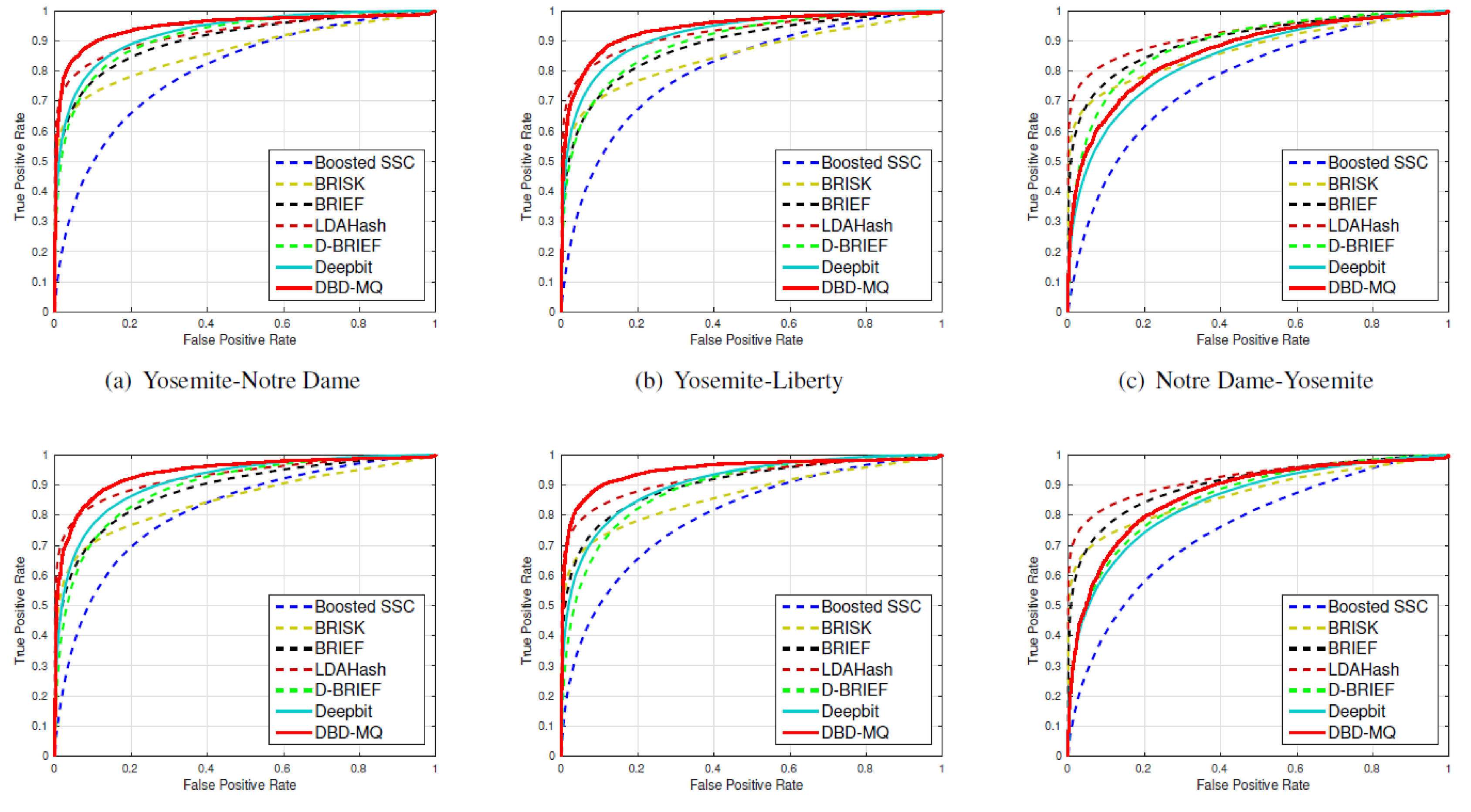
Experiments

- The CIFAR-10 dataset



(a) 16 bits (b) 32 bits (c) 64 bits

- The Brown dataset



(a) Yosemite-Notre Dame (b) Yosemite-Liberty (c) Notre Dame-Yosemite

(d) Notre Dame-Liberty (e) Liberty-Notre Dame (f) Liberty-Yosemite

Train Test	Yosemite Noter Dame	Yosemite Liberty	Notre Dame Yosemite	Notre Dame Liberty	Liberty Notre Dame	Liberty Yosemite	Average ERR
SIFT (128 bytes)	28.09	36.27	29.15	36.27	28.09	29.15	31.17
Boosted SSC (16 bytes)	72.20	71.59	76.00	70.35	72.95	77.99	73.51
BRISK (64 bytes)	74.88	79.36	73.21	79.36	74.88	73.21	75.81
BRIEF (32 bytes)	54.57	59.15	**54.96**	59.15	54.57	**54.96**	56.23
DeepBit (32 bytes)	29.60	34.41	63.68	32.06	26.66	57.61	40.67
LDAHash (16 bytes)	51.58	49.66	52.95	49.66	51.58	52.95	51.40
D-BRIEF (4 bytes)	43.96	53.39	46.22	51.30	43.10	47.29	47.54
BinBoost (8 bytes)	14.54	21.67	18.96	20.49	16.90	22.88	19.24
RFD (50-70 bytes)	11.68	19.40	14.50	19.35	13.23	16.99	15.86
DBD-MQ (32 bytes)	**27.20**	**33.11**	57.24	**31.10**	**25.78**	57.15	**38.59**

Email: duanyq14@mails.tsinghua.edu.cn